International Journal of Forecasting 31 (2015) 325–348

Contents lists available at ScienceDirect

International Journal of Forecasting

journal homepage: www.elsevier.com/locate/ijforecast

Macroeconomic information, structural change, and the prediction of fiscal aggregates

Andrea Carriero [a,*], Haroon Mumtaz [a], Angeliki Theophilopoulou [b]

[a] *Queen Mary, University of London, United Kingdom*
[b] *Westminster Business School, United Kingdom*

ARTICLE INFO

Keywords:
Bayesian VARs
Forecasting
Fiscal policy

ABSTRACT

Previous research on the prediction of fiscal aggregates has shown evidence that simple autoregressive models often provide better forecasts of fiscal variables than multivariate specifications. We argue that the multivariate models considered by previous studies are small-scale, probably burdened by overparameterization, and not robust to structural changes. Bayesian Vector Autoregressions (BVARs), on the other hand, allow the information contained in a large data set to be summarized efficiently, and can also allow for time variation in both the coefficients and the volatilities. In this paper we explore the performance of BVARs with constant and drifting coefficients for forecasting key fiscal variables such as government revenues, expenditures, and interest payments on the outstanding debt. We focus on both point and density forecasting, as assessments of a country's fiscal stability and overall credit risk should typically be based on the specification of a whole probability distribution for the future state of the economy. Using data from the US and the largest European countries, we show that both the adoption of a large system and the introduction of time variation help in forecasting, with the former playing a relatively more important role in point forecasting, and the latter being more important for density forecasting.

1. Introduction

The forecasting of future developments in fiscal variables has been of increasing importance in recent years, especially since the latest financial and Euro-area sovereign debt crisis. Fiscal positions affect the credit risk and the sovereign cost of borrowing. Following the latest developments in the international debt markets, it is apparent that markets take economic fundamentals into account seriously, and penalize countries heavily for fiscal imbalances (von Hagen, Schuknecht, & Wolswijk, 2011). Arghyrou and Kontonikas (2012) claim that there has been a significant shift in market behavior since 2007, from a convergence-based pricing model to a fundamentals-based pricing model, meaning that forecasting fundamental macroeconomic variables has become more important.

A number of institutions produce forecasts of the macroeconomic variables that provide important feedback for their policies. For example, central banks need to identify the impact of fiscal policies on fundamentals and inflation in order to conduct monetary policy. OECD and IMF use forecasts to determine whether they need to intervene, and to provide recommendations to individual countries about the sustainability of their fiscal and monetary policies. National research institutes and rating agencies, which have to assess the default risk entailed in the debt

* Corresponding author.
E-mail addresses: a.carriero@qmul.ac.uk (A. Carriero), h.mumtaz@qmul.ac.uk (H. Mumtaz), a.theophilopoulou@westminster.ac.uk (A. Theophilopoulou).

http://dx.doi.org/10.1016/j.ijforecast.2014.06.006

securities issued by governments, use forecasts to express their own views on the fiscal and monetary policies followed by national authorities.

Forecasts are also of increasing importance in the determination of budgetary goals. For instance, if high economic growth is forecasted, a government would expect a higher level of structural revenues, and can budget a higher discretionary expenditure. However, official growth forecasts for the Euro-area have been systematically over-optimistic (see Artis & Marcellino, 2001; Jonung & Larch, 2006; Strauch, Hallerberg, & von Hagen, 2004), while the evidence for the US is mixed (Leal, Pérez, Tujula, & Vidal, 2008). Optimistic forecasts from official authorities may have played an important role in the excessive deficits which have been observed in several Euro countries (Jonung & Larch, 2006). Possible reasons for the observed bias may be that governments and official authorities may have non-symmetric loss functions, or that significant economic variables have been omitted from the estimated models. Another type of error may also be introduced via miscalculated fiscal variables, the estimation of which is based on economic variables such as the output gap and GDP volatility (Cassidy, Kamlet, & Nagin, 1989; Feenberg, Gentry, Gilroy, & Rosen, 1989; Leal et al., 2008; Melander, Sismanidis, & Grenouilleau, 2007).

Favero and Marcellino (2005) provide a comprehensive study on the forecasting of fiscal variables using a wide range of econometric models. In particular, they consider univariate autoregressive and moving average models, vector autoregressions (VARs), and small-scale semistructural models, and compare them with institutional forecasts made by the OECD. Their results show that simple time series univariate methods work well and are able to deliver unbiased forecasts, or slightly upward-biased forecasts for the debt–GDP dynamics, whereas the OECD forecasts are typically biased. The fact that univariate models work better than multivariate ones is puzzling, because economic theory would suggest that fiscal variables should be tightly intertwined, and therefore, in theory, models based on a system of macroeconomic variables, such as VARs, should produce better forecasts than simple univariate specifications.

Favero and Marcellino (2005) suggest that these results are due mostly to the short sample available with respect to the number of coefficients to be estimated (overparameterization), the robustness of simple methods to structural breaks, and the difficulty of modelling the joint behaviors of several variables in a period of substantial institutional and economic change.

In this paper, we consider using econometric models that can deal with these problems efficiently. The models that we consider allow the information contained in a large data set to be used without incurring the overparameterization problem, and can allow for time variation in the coefficients and the volatilities, a characteristic which makes them robust to structural changes, regardless of whether such changes happen smoothly or abruptly (breaks). We show that once overparameterization and structural change have been dealt with appropriately, multivariate models do provide a better description of the macroeconomy than univariate specifications, and the use of a large panel of macroeconomic data does provide improvements in terms of forecast accuracy.

A key aspect in the forecasting of macroeconomic variables is the assessment of the overall uncertainty that exists around point forecasts. This aspect seems particularly relevant for fiscal variables, as assessments of fiscal stability and of the overall credit and default risk of a country should typically be based on the specification of a complete probability distribution for the future state of the economy.

Accordingly, the second contribution of this paper is to focus on forecasting the whole predictive distribution of fiscal variables, rather than limiting the interest to point forecasts only. Our empirical results show that the use of models that allow for drifting coefficients and volatility does provide a better characterization of the uncertainty in the economy, which translates into substantial gains in density forecasts with respect to simpler specifications with constant coefficients and volatilities.

To deal with the problem of overparameterization, we consider the use of Bayesian Vector Autoregressions (BVARs). BVARs have a long history in forecasting, stimulated by their effectiveness, as documented in the seminal studies of Doan, Litterman, and Sims (1984) and Litterman (1986). In recent years, these models seem to have been being used even more systematically for policy analysis and the forecasting of macroeconomic variables. Starting from the paper of Banbura, Giannone, and Reichlin (2010), the benefits of using BVARs for macroeconomic forecasting using large data-sets have been documented in several recent papers (e.g., Carriero, Clark, & Marcellino, 2013a; Carriero, Kapetanios, & Marcellino, 2011; Koop, 2013). The good performance of BVARs is not limited to large data sets with large numbers of parameters to be estimated: Litterman (1986) has shown that in average-sized models with up to six variables, forecasts of BVARs without judgemental adjustments are at least as good and competitive as the best commercially available forecasts.

To account for the possibility that the data generating processes of fiscal variables have experienced changes in behavior over time, we estimate time-varying parameter Bayesian vector autoregressions (TVP-BVARs). We allow for both the autoregressive coefficients and the variance of the errors varying over time. The TVP-BVAR model has also been used in other studies such as that of D'Agostino, Gambetti, and Giannone (2013), who find that it outperforms other methods such as a fixed coefficient VAR and a time varying AR for the forecasting of US unemployment, inflation and short term interest rates. Clark (2011) finds similar results for a model featuring only time variation in volatility.

We explore the performances of BVARs with constant and drifting coefficients in forecasting key fiscal variables such as government revenues, expenditures, and interest payments on the outstanding debt. We focus on forecasting fiscal variables for the US and the three largest E.U. economies, namely the UK, France, and Germany. These European countries were all involved in the European sovereign debt crisis either directly or indirectly, and they still face enormous fiscal constraints with relevant economic implications. The inclusion of European countries that were affected by the recent sovereign debt crisis more directly, such as Italy, Spain, and Greece, would be desirable; however, the data available for these countries are either too short or too unreliable to allow the estimation of time varying parameter specifications, and therefore we have not included them in the comparison.

We obtain two main results from our empirical analysis. First, we find that all BVAR specifications (with fixed or time-varying coefficients) outperform simple autoregressive forecasts overall, while the forecasts produced by a classical VAR do not. This confirms that classical VARs are likely to be burdened by overparameterization, but that, once overparameterization has been dealt with, the use of additional explanatory variables does help in forecasting fiscal variables, and multivariate models should be preferred to univariate specifications. Second, we find that both the adoption of a large system and the introduction of time variation help in forecasting, with the former playing a relatively more important role in point forecasting, and the latter being more important for density forecasting.

The paper is structured as follows: Section 2 describes the two main Bayesian models with constant and time-varying coefficients. Section 3 describes the data set and the forecasting scheme. Section 4 provides the results of our empirical application. Section 5 concludes. Two appendices provide details on the BVAR and TVP-BVAR models.

2. Models

Vector autoregressions (VARs) have been used widely in macroeconomics to study the interactions between fiscal policy and other macroeconomic variables (see e.g. Blanchard & Perotti, 2002, for the US; and Marcellino, 2006, for the largest countries in the Euro area). The advantage of VARs lies in their rather general representation, which does not require any restrictions to be imposed on the parameters, thus enabling the models to capture complex relationships among variables.

However, the large numbers of parameters that need to be estimated reduce the degrees of freedom, leading to less accurate estimates, a situation that is usually referred to as the "curse of dimensionality". Moreover, classical VARs do not allow for drifts in coefficients and volatilities, while several recent studies, such as those of Carriero, Clark, and Marcellino (2012), Clark (2011), Cogley and Sargent (2005), D'Agostino et al. (2013), Koop and Korobilis (2013), and Primiceri (2005), have emphasized that the inclusion of drifting coefficients and volatility is key for understanding the dynamics of macroeconomic variables, as well as for forecasting.

A Bayesian vector autoregression (BVAR) is a VAR with coefficients that are random variables on which the researcher can impose some a priori information. The content of such prior information can be based on both judgment and a pre-sample of data. The precision of such prior information can vary from very high to very loose, a choice which is left entirely up to the researcher.

Without entering into philosophical disputes about the Bayesian and classical approaches to econometrics, we think that it is worth stressing here that BVARs can be interpreted simply as a selection device. Consider the first equation of a large VAR: there are many regressors, and the researcher needs to solve the trade-off between using as much information as possible, and the loss in degrees of freedom that comes from having too many parameters to estimate. The researcher can follow the simple method of having a set of variables, adding one variable at a time, testing for its significance, and then accepting or rejecting it. Implicitly, what this procedure does is select the regressors on the basis of how much valuable information they contain. Information is valuable if it is able to increase the likelihood of the model significantly. However, in this case, the order in which the various candidate variables are considered for inclusion in the model can influence the outcome crucially.

The Bayesian approach works similarly, in the sense that it also weights each coefficient in the model relative to the contribution that the regressor to which that coefficient is attached makes to the likelihood of the model. The coefficient attached to a given candidate regressor is set to some prior value (for example 0), and only if the information contained in the data is valuable enough to influence the likelihood will the posterior mean of the coefficient be far from its prior value. More precisely, rather than acting as a selection device which either includes or excludes a regressor, the BVAR includes all of the regressors, but assigns a different weight to each of them. The weight increases with the informational content of a given regressor, i.e. the higher its contribution to the likelihood, the higher its weight.

Moreover, BVARs can accommodate time variation in the coefficients easily. As BVAR coefficients are random variables, it is natural to model them as stochastic processes and treat them as unobserved variables which can be filtered out easily using an MCMC algorithm.

In what follows, we describe the two models we use in this paper, namely the constant coefficient BVAR with natural conjugate prior, and the time varying parameters BVAR (TVP-BVAR).

2.1. BVARs with constant coefficients (BVARs)

We use a standard BVAR with a Normal-Inverted Wishart (N-IW) natural conjugate prior. Given N different variables grouped in the vector $y_t = (y_{1t}\ y_{2t}\ \ldots\ y_{Nt})'$, we consider the following vector autoregression (VAR):

$$y_t = \Phi_c + \Phi_1 y_{t-1} + \Phi_2 y_{t-2} + \cdots + \Phi_p y_{t-p} + \Sigma^{1/2}\varepsilon_t;$$
$$\varepsilon_t \sim i.i.d.\ N(0, I_N), \tag{1}$$

where $t = 1, \ldots, T$. Each equation has $M = Np + 1$ regressors. By grouping the coefficient matrices in the $N \times M$ matrix $\Phi = [\Phi_c\ \Phi_1\ \ldots\ \Phi_p]$ and defining $x_t = (1\ y'_{t-1}\ \ldots\ y'_{t-p})'$ as a vector containing an intercept and p lags of y_t, the VAR in Eq. (1) can be written as:

$$y_t = \Phi x_t + \varepsilon_t.$$

Considering all data points $t = 1, \ldots, T$ gives:

$$Y = X\Phi + E,$$

where $Y = [y_1, \ldots, y_T]'$, $X = [x_1, \ldots, x_T]'$, and $E = [\varepsilon_1, \ldots, \varepsilon_T]'$ are, respectively, $T \times N$, $T \times M$ and $T \times N$ matrices.

For the VAR coefficients and error variances, we use the conjugate N-IW prior:

$$\Phi|\Sigma \sim N(\Phi_0, \Sigma \otimes \Omega_0), \qquad \Sigma \sim IW(S_0, v_0).$$

As the N-IW prior is conjugate, the conditional posterior distribution of this model is also N-IW (Zellner, 1971):

$$\Phi|\Sigma, Y \sim N(\bar{\Phi}, \Sigma \otimes \bar{\Omega}), \qquad \Sigma|Y \sim IW(\bar{S}, \bar{v}), \tag{2}$$

where $\bar{\Phi} = \bar{\Omega}(\Omega_0^{-1}\Phi_0 + X'Y)$, $\bar{\Omega}^{-1} = \Omega_0^{-1} + X'X$, $\bar{v} = v_0 + T$, and $\bar{S} = S_0 + Y'Y + \Phi_0'\Omega_0^{-1}\Phi_0 - \bar{\Phi}'\bar{\Omega}^{-1}\bar{\Phi}$. In the case of the natural conjugate N-IW prior, the marginal posterior distribution of Φ is matricvariate-t with expected value $\bar{\Phi}$.

In our baseline specification, we impose the prior expectation and variance of the coefficient matrices to be:

$$\mathrm{E}[\Phi_k^{(ij)}] = \begin{cases} \Phi^* & \text{if } i = j,\ k = 1 \\ 0 & \text{otherwise,} \end{cases}$$

$$\mathrm{Var}[\Phi_k^{(ij)}] = \theta \frac{1}{k^2}\sigma_i^2/\sigma_j^2, \quad k = 1, \ldots, p, \tag{3}$$

where $\Phi_k^{(ij)}$ denotes the element in position (i, j) in the matrix Φ_k, and where the covariances among the coefficients in Φ_k are zero. As will be discussed in Section 3.1, all series are transformed to stationarity before estimation is performed, so we set $\Phi^* = 0$.[1] The shrinkage parameter θ measures the tightness of the prior: when $\theta \to 0$, the prior is imposed exactly and the data do not influence the estimates, while as $\theta \to \infty$, the prior becomes loose and the prior information does not influence the estimates, which will approach the standard *OLS* estimates. We will discuss the choice of this parameter in detail below. The factor σ_i^2/σ_j^2 is a scaling parameter which accounts for the different scale and variability of the data. To set the scale parameters σ_i^2, we follow common practice (see e.g. Litterman, 1986; Sims & Zha, 1998) and set them equal to the variance of the residuals from a univariate autoregressive model for the variables.

The prior specification is completed by choosing v_0 and S_0 so that the prior expectation of Σ is equal to a fixed diagonal residual variance $\mathrm{E}[\Sigma] = diag(\sigma_1^2, \ldots, \sigma_N^2)$. In particular, following Kadiyala and Karlsson (1997), we set the diagonal elements of S_0 to $s_{0ii} = (v_0 - N - 1)\sigma_i^2$ and $v_0 = N + 2$.

The prior described above is similar to that proposed by Sims and Zha (1998), with the subtle difference that, in the original implementation, the prior is elicited on the coefficients of the structural representation of the VAR rather than on the reduced form. This prior has been used widely in the literature, see e.g. Leeper, Sims, and Zha (1996), Robertson and Tallman (1999), Waggoner and Zha (1999), and Zha (1998), and more recently Giannone, Lenza, and Primiceri (2012).

To make the prior operational, one needs to choose the value of the hyperparameter θ, which controls the tightness of the prior. We follow Carriero et al. (2013a) and choose θ at each point by maximizing the marginal data density of the model:[2]

$$\theta_t^* = \arg\max_{\theta} \ln p(Y). \tag{4}$$

The marginal data density can be obtained by integrating out all of the coefficients in the model. Defining Θ as the set of all coefficients in the model, we have:

$$p(Y) = \int p(Y|\Theta)p(\Theta)d\Theta.$$

Under the N-IW prior, the density $p(Y)$ can be computed in closed form, and is given in Appendix A. We optimize over a discrete grid $\theta \in \{0.01, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 1, 10\}$.[3] The marginal likelihood is also used to select the lag length p of the system:

$$p^* = \arg\max_{p} \ln p(Y), \tag{5}$$

where we optimize over the grid $p = 1, 2, \ldots, 12$.

Under the standard N-IW prior described above, the full distribution of the 1-step ahead forecasts is given by:

$$y'_{T+1}|x'_{T+1} \sim MT(x'_{T+1}\bar{\Phi}, (x'_{T+1}\bar{\Omega}x_{T+1} + 1)^{-1}, \bar{S}, \bar{v}).$$

Multi-step-ahead forecasts obtained by iteration are not available in closed form, but can be simulated using a MC algorithm which draws a sequence of (Σ, Φ) using Eq. (2),[4] with shocks $\varepsilon_{t+1}, \ldots, \varepsilon_{t+h}$, and computes the implied path of $\hat{y}_{t+h}^{(j)}$ at each draw j. We form point forecasts as means of the draws of simulated forecasts.

2.2. BVARs with time varying coefficients (TVP-BVARs)

We use the specification of D'Agostino et al. (2013):

$$y_t = \Phi_{c,t} + \Phi_{1,t}y_{t-1} + \Phi_{2,t}y_{t-2} + \cdots + \Phi_{p,t}y_{t-p} + \Sigma_t^{1/2}\varepsilon_t; \quad \varepsilon_t \sim i.i.d.\ N(0, I_N), \tag{6}$$

where y_t denotes the variables in the VAR. Note that Eq. (6) is more general than Eq. (1), as it allows for variation

[1] In the traditional implementation with data in levels, Φ^* is set to one to reflect the idea that the variables in the VAR follow univariate random walks. The specification $\Phi^* = 0$ on the differenced data is equivalent, as it again reflects the belief that the levels of the data follow univariate random walks.

[2] Giannone et al. (2012) propose a similar strategy for forecasting a macroeconomic dataset. While our strategy implicitly assumes a flat prior on a discrete set of possible values for θ, their strategy assumes a proper (albeit uninformative) prior on a continuum of values. A second alternative would be to calibrate θ so that the average in-sample fit for the fiscal variables is the same as in the univariate model we use as a benchmark (described in Section 3.3). This latter strategy was used by Banbura et al. (2010), and can be treated as a robustness check.

[3] It turns out that the value of the tightness does not change substantially over time. For example, with US data, the optimal θ is equal to 0.4 for most of the sample, with only a few exceptions towards the end of the sample, where it sometimes decreases to 0.3.

[4] Drawing a sequence of Φ can be computationally demanding in general, but in this specific case the matricvariate structure of the N-IW prior ensures the existence of a factorization that speeds up the computations considerably. Indeed, letting V be a $M \times N$ matrix drawn from a matricvariate standard normal distribution, we can draw the matrix Φ as follows:

$$\Phi = \bar{\Phi} + chol(\bar{\Omega}) \times V \times chol(\Sigma)'.$$

This can speed up the computations considerably, because the two Cholesky decompositions $chol(\bar{\Omega})$ and $chol(\Sigma)$ require only $M^3 + N^3$ operations.

in the coefficients ($\Phi_{\cdot,t}$) and in the volatilities ($\Sigma_t^{1/2}$). The VAR coefficients $\Phi_t = \{\Phi_{c,t}, \Phi_{j,t}\}$ evolve as random walks

$$\Phi_t = \Phi_{t-1} + \eta_t.$$

Following Cogley and Sargent (2005), the covariance matrix of the innovations $v_t = \Sigma_t^{1/2}\varepsilon_t$ is factored as:

$$\text{VAR}\,(v_t) \equiv \Sigma_t = A_t^{-1} H_t (A_t^{-1})'.$$

The time-varying matrices H_t and A_t are defined as:

$$H_t = \begin{bmatrix} h_{1,t} & 0 & 0 & 0 \\ 0 & h_{2,t} & 0 & 0 \\ 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & h_{N,t} \end{bmatrix}; \qquad A_t = \begin{bmatrix} 1 & 0 & 0 & 0 \\ \alpha_{2,1,t} & 1 & 0 & 0 \\ \dots & \dots & 1 & 0 \\ \alpha_{N,1,t} & \dots & \alpha_{N,N-1,t} & 1 \end{bmatrix}, \tag{7}$$

with the $h_{i,t}$ evolving as geometric random walks:

$$\ln h_{i,t} = \ln h_{i,t-1} + \tilde{\nu}_t. \tag{8}$$

Following Primiceri (2005), we postulate the non-zero and non-one elements of the matrix A_t to evolve as driftless random walks,

$$\alpha_t = \alpha_{t-1} + \tau_t,$$

and we assume the vector $[v_t', \eta_t', \tau_t', \tilde{\nu}_t']'$ to be distributed as an i.i.d. multivariate normal with mean 0 and variance V:

$$V = \begin{bmatrix} \Sigma_t & 0 & 0 & 0 \\ 0 & Q & 0 & 0 \\ 0 & 0 & S & 0 \\ 0 & 0 & 0 & G \end{bmatrix},$$

where G is a diagonal matrix with diagonal elements $\sigma_1^2, \dots, \sigma_N^2$.

The prior distributions and the conditional posteriors are described in Appendix B. We point out two aspects here. First, the prior for Q is set using a pre-sample of T_0 quarters. In particular, let Q_{OLS} denote the OLS estimate of the coefficient covariance matrix using the training sample. The prior distribution for Q is assumed to be Inverse Wishart, with the scale matrix given by $\bar{Q} = Q_{OLS} \times T_0 \times k$, where the scalar $k = 3.5 \times 10^{-4}$, as per Cogley and Sargent (2005). The prior degrees of freedom are set equal to T_0, i.e., the length of the training sample.

The model can be estimated using a Gibbs sampling algorithm, which is described in more detail in Appendix B. Following Cogley and Sargent (2005), we require the VAR coefficients to be stable at each point in time, and achieve this via rejection sampling. Note that finding stable draws of the VAR coefficients for large systems can be very time-consuming, and the problem is magnified in our application because the model is estimated recursively. We therefore limit our consideration to systems with at most four variables.

The lag length is set to two for the US and the UK, reflecting convention in previous TVP-BVAR studies for these countries (see Cogley & Sargent, 2005, for the US, and Cogley, Morozov, & Sargent, 2005, for the UK). For the remaining countries, the lag length is set to one, because the time series available are severely limited.[5]

For the production of forecasts, we follow D'Agostino et al. (2013) and fix the future values of the time-varying matrix Φ_t at its estimated value. In particular, given the value of the coefficient matrix Φ_t at time t, the forecasts are produced by using Eq. (6) in periods $t+1, \dots, t+h$, where the right-hand side variables are either the actual data (when available for a particular forecast horizon) or the forecasts produced for the previous period. The volatilities $h_{i,t}$ are instead simulated forward using Eq. (8). We form point forecasts as means of the draws of simulated forecasts.

Finally, in order to ascertain the roles of the time variation in the coefficients and in the volatilities separately, we will consider a special case of this model in which the variation in the conditional mean parameters Φ_t is removed, so that these parameters are constant ($\Phi_t = \Phi$), while the volatility matrix $\Sigma_t^{1/2}$ still changes over time. This model is in between the TVP-BVAR and the BVAR, because with respect to the former it does not feature change in the conditional mean parameters, while with respect to the latter it does feature variation in the volatilities. We label this model SV-BVAR, where SV stands for stochastic volatility.

3. Empirical application

3.1. Data

We use the above forecasting methods to forecast fiscal aggregates of four large economies: the US, the UK, France, and Germany. The sample is at a quarterly frequency, and its length is different for each country, mainly because of data availability. The data for the US span the period from 1969:Q1 to 2010:Q4. For the UK, the time span is 1972:Q1–2010:Q4. For France and Germany, the data run from 1991:Q2 to 2010:Q4 and from 1991:Q1 to 2010:Q3, respectively. As the sample sizes for France and Germany are particularly small, we anticipate here that a good deal of caution will be needed in interpreting the results for these two countries, especially with regard to the time-varying coefficient specifications.[6]

For each country, we consider three fiscal and six additional macroeconomic variables which we use to forecast the fiscal variables. The set of fiscal variables contains government expenditure, interest payments on public debt, and tax revenues. All of these variables are taken

[5] There are two obstacles to increasing the number of lags in the TVP specifications. First, an overly rich dynamic specification would encounter serious computational issues when drawing the vector of coefficients, because the dimension of the parameter vector would imply an extremely high percentage of draws in which at least some roots lie outside the unit circle. This is a problem that has notoriously affected TVP-VARs in general, and is very difficult to circumvent; see for example the discussions by Koop (2013) and Koop and Korobilis (2013). The second problem is specific to our application, and in particular to the fact that some of the countries in our exercise have very short data sets in the time series dimension. While this is not a problem for the constant coefficient BVARs, it is problematic for the TVP-VARs.

[6] Note that the first 40 (20) observations are used as a training sample when estimating the TVP-VAR for the US and the UK (Germany and France).

Table 1
Variable transformations.

Variable, $x_{i,t}$	Transformation, $y_{i,t}$
Government expenditures, ratio to GDP	$y_{i,t} = \Delta \log(x_{i,t})$
Tax revenues, ratio to GDP	$y_{i,t} = \Delta \log(x_{i,t})$
Interest payments on public debt, ratio to GDP	$y_{i,t} = \Delta \log(x_{i,t})$
GDP growth	$y_{i,t} = x_{i,t}$
Short term interest rate	$y_{i,t} = \Delta(x_{i,t})$
Long term interest rate	$y_{i,t} = \Delta(x_{i,t})$
Inflation	$y_{i,t} = x_{i,t}$
Consumption	$y_{i,t} = \Delta \log(x_{i,t})$
Industrial production	$y_{i,t} = \Delta \log(x_{i,t})$

Data for the US are taken from the FRED database (Federal Reserve Bank of St. Louis), data for the European countries are taken from Eurostat, Bloomberg and Bundesbank.

as ratios to GDP. We do not include the deficit (either primary or with interest) in the analysis, as, broadly speaking, it is a linear combination of the three fiscal variables we consider, and therefore its inclusion might cause issues of collinearity. A similar argument applies to the level of public debt, as such a variable would need to be included in its first difference in order to achieve stationarity, and the resulting variable would be highly collinear with the deficit.

It is worth noting that, for Germany and France, the Stability and Growth pact and the Fiscal Compact impose constraints on some fiscal ratios. This is an interesting fact that could potentially be taken into account in the models and used to forecast fiscal ratios better. However, these constraints were violated on some occasions, so it is not clear how binding they are. Moreover, fiscal forecasts are typically used by institutions (such as the OECD and the IMF) to assess whether they need to intervene and to provide recommendations to individual countries about the sustainability of their fiscal policies. Due to these considerations, having a model which can produce forecasts that violate such constraints is desirable.

The remaining macroeconomic variables are GDP growth, consumption, inflation, industrial production, and short and long term interest rates. We proxy short term rates by taking three-month yields on Treasury Bills, and long term rates by ten-year yields on government bonds. The GDP is the gross domestic product at market prices. To account for inflation, we use a price index with base year 2000. For industrial production (IP), we have used an index with base year 2000, and all industries are included. Consumption is the final consumption expenditure as a ratio to GDP.

Some of the information contained in the IP and the consumption expenditure might be reflected in the GDP already. However, there is a rationale for using some disaggregation in the data. For example, IP provides specific information on the production of goods, while changes in consumption might reflect the fact that consumers are (usually correctly) reflecting their forecasts of a weakening or strengthening job market (e.g. Breeden, 2012). Several papers have provided evidence that a certain level of disaggregate information can provide additional gains in forecasting, even at a quarterly frequency, see e.g. Marcellino, Stock, and Watson (2006) and Stock and Watson (2002).

We did not include unemployment in the pool of regressors because this variable had substantial differences in definition, construction and length between countries. Moreover, it is typically considered to be a lagging variable with respect to the cycle, which is likely to make it less useful as a forecasting tool.[7]

We have used a number of data sources. All data for the US economy come from the database of the Federal Reserve Bank of St. Louis. Most fiscal and other macroeconomic data for the European countries are from Eurostat. Quarterly data for three-month Treasury bills and ten-year government bonds with constant maturity for Germany and France are collected from Bloomberg. The time series for industrial production and final consumption expenditure for Germany come from the Bundesbank.

All variables are seasonally adjusted when necessary. All of the models are estimated after transforming the variables as necessary to obtain stationarity. However, the forecasting results we provide are based instead on the original variables.[8] The transformations used on each variable are listed in Table 1.

3.2. Forecasting exercise

The forecasting exercise is performed in pseudo real time, i.e., we never use information which is not available at the time the forecast is made, but do use final vintage data. For all models, we use a recursive estimation scheme. For example, for the US, the initial estimation window ranges from 1969:Q1 to 1989:Q1, which includes 20 years (i.e., 80 quarterly data-points). Forecasts up to four steps ahead are produced to cover the period 1989:Q2 to 1990:Q1. Then, the estimation window is augmented with an additional data-point, becoming 1969:Q1 to 1989:Q2, and forecasts up to 1991:Q2 are produced. This scheme continues until the last estimation window, which is 1969:Q1 to 2009:Q4, and yields forecasts for the period 2010:Q1 to 2010:Q4. The forecast evaluation window is therefore from 1989:Q2 to 2010:Q4. For the remaining countries, the forecast evaluation windows are: 1992:Q2 to 2010:Q4 for the UK, 2001:Q3 to 2010:Q4 for France, and 2001:Q2 to 2010:Q3 for Germany.

7 The lagging nature of unemployment is particularly marked for the European countries in a sample that includes the recent crisis.

8 This means that we produce the forecasts of the transformed variable and then get the forecast of the original variable by inverting the transformation.

3.3. Forecast evaluation

We evaluate both point and density forecasts of the models examined. As a benchmark model, we will use the following Bayesian autoregressive model (BAR):

$$y_t^{(i)} = \phi_c + \phi_1 y_{t-1}^{(i)} + \cdots + \phi_p y_{t-p}^{(i)} + \sqrt{\sigma_i^2}\varepsilon_t^{(i)}; \quad \varepsilon_t \sim i.i.d.\ N(0,1). \tag{9}$$

Favero and Marcellino (2005) provide evidence that univariate models such as the BAR can produce very good point forecasts of fiscal variables for several European countries.

Note that the model in Eq. (9) can be obtained as a special case of the BVAR in Eq. (1), just by setting the number of variables to $N = 1$. Therefore, the estimation of the benchmark model and the production of the forecasts proceed precisely as in the BVAR. In particular, we specify the same prior mean and variances as in the BVAR, and the resulting posteriors are the univariate versions of Eq. (2). Similarly, we choose the overall tightness and lag-length via the marginal likelihood using Eqs. (4) and (5) respectively. Finally, the forecasts are produced using an MC algorithm which draws a sequence of $\sigma_i, \phi_c, \phi_1, \ldots, \phi_p$ (using the univariate versions of Eq. (2)) and the shocks $\varepsilon_{t+1}^{(i)}, \ldots, \varepsilon_{t+h}^{(i)}$, and computes the implied path of $\hat{y}_{t+h}^{(j)}$ at each draw j. Point forecasts are computed as means of the draws of simulated forecasts. For robustness, we also computed results for a classical autoregressive model with the lag length selected via the BIC; the results were similar to those obtained with the BAR, and therefore we do not provide them here, to save space.

Typically, stochastic volatility improves density forecasts, and hence it is helpful to include as a benchmark a version of the BAR in which the volatilities are drifting. This model can also be obtained as a special case of one of our more general models, the TVP-BVAR in Eq. (6), simply by setting $N = 1$ and removing the variation in the conditional mean coefficients Φ_t. The model is:

$$y_t^{(i)} = \phi_c + \phi_1 y_{t-1}^{(i)} + \cdots + \phi_p y_{t-p}^{(i)} + \sqrt{\sigma_{i,t}^2}\varepsilon_t^{(i)}; \quad \varepsilon_t \sim i.i.d.\ N(0,1) \tag{10}$$

$$\ln \sigma_{i,t}^2 = \ln \sigma_{i,t}^2 + u_t; \quad u_t \sim i.i.d.\ N(0, \omega_i^2). \tag{11}$$

Estimates and forecasts for this model are produced using the same algorithm as the TVP-BVAR, modified appropriately to take into account the constancy of the coefficients Φ_t and the univariate dimension of the problem. We label the benchmark model in Eqs. (10)–(11) the SV-BAR.

Finally, to provide a general indication of the overall forecastability of the series under analysis, we also add to the pool of models a simple random walk no-change forecast, which we label RW. Such a forecast is obtained by simulating forward the equation $y_t^{(i)} = y_{t-1}^{(i)} + \sqrt{\sigma_i^2}\varepsilon_t^{(i)}$. The coefficient σ_i is estimated using the sample variance of $y_t^{(i)} - y_{t-1}^{(i)}$, and forecasts are obtained by drawing the shocks $\varepsilon_{t+1}^{(i)}, \ldots, \varepsilon_{t+h}^{(i)}$ and multiplying them by the estimated $\sqrt{\sigma_i^2}$.

3.3.1. Point forecasts

For point forecasts, we evaluate our results in terms of the root mean squared forecast error (RMSFE) for a given model. Let $\hat{y}_{t+h}^{(i)}(M)$ denote the forecast of the ith variable $y_{t+h}^{(i)}$ made by model M. The RMSFE of model M in forecasting the ith variable at horizon h is:

$$\mathrm{RMSFE}_{i,h}^{M} = \sqrt{\frac{1}{P}\sum\left(\hat{y}_{t+h}^{(i)}(M) - y_{t+h}^{(i)}\right)^2},$$

where the sum is computed over all P forecasts produced. Results are reported in terms of the relative RMSFE (RelRMSFE) between the model and the benchmark:

$$\mathrm{RelRMSFE}_{i,h}^{M} = \frac{\mathrm{RMSFE}_{i,h}^{M}}{\mathrm{RMSFE}_{i,h}^{\mathrm{BAR}(p^*)}}. \tag{12}$$

A RelRMSFE below one signals that model M outperforms the benchmark.

To provide a rough gauge as to whether the RelRMSFE ratios are significantly different from one, we use the Diebold and Mariano (1995) t-statistic for equal MSFEs, applied to the forecast of each model relative to the benchmark. Our use of the Diebold–Mariano test with forecasts that are sometimes nested is a deliberate choice. Monte Carlo evidence from Clark and McCracken (2013) indicates that, with nested models, the Diebold–Mariano test can be viewed as a somewhat conservative (conservative in the sense of tending to have a size that is modestly below the nominal size) test for equal accuracy in a finite sample, relative to normal critical values. Since our proposed model can be viewed as nesting the benchmarks we are using for comparison, we treat the tests as one-sided, and only reject the benchmark in favor of the null (i.e., we do not consider rejections of the alternative model in favor of the benchmark). The underlying p-values are based on t-statistics computed with a serial correlation-robust variance, using a rectangular kernel, $h - 1$ lags, and the small-sample adjustment of Harvey, Leybourne, and Newbold (1997).

Finally, to provide an overall measure of the forecastability of the models, we have also considered the mean absolute percentage error (MAPE):

$$\mathrm{MAPE}_{i,h}^{M} = \frac{1}{P}\sum\left|\frac{\hat{y}_{t+h}^{(i)}(M) - y_{t+h}^{(i)}}{y_{t+h}^{(i)}}\right|,$$

where the sum is computed over all the P forecasts produced. To avoid cluttering up the tables, we only comment on the values of the MAPE achieved by some selected forecasting models.

3.3.2. Density forecasts

The overall calibration of the density forecasts can be measured using the average of log predictive likelihoods (density scores henceforth), motivated and described by, e.g., Geweke and Amisano (2010). For model M_i, the h-step-ahead score is defined as:

$$\mathrm{SCORE}_{i,h}^{M} = \frac{1}{P}\sum \log p(y_{t+h}^{(i)} | y^{(t)}, M),$$

where the sum is computed over all the P forecasts produced, $y_{t+h}^{(i)}$ denotes the observed outcome for the data in period $t+h$, and $y^{(t)}$ denotes the history of data up to period t (the sample used to estimate the model and form the prediction for period $t+h$). The predictive density $p(\cdot)$ is obtained by univariate kernel estimation based on the MCMC output. Again, we compare this model against the BAR(p^*) benchmark given by Eq. (9). In this case, the SCORE is a gain function, and in logarithmic scale, and therefore we consider its differences:

$$\Delta\,\mathrm{SCORE}_{i,h}^{M} = \mathrm{SCORE}_{i,h}^{M} - \mathrm{SCORE}_{i,h}^{\mathrm{BAR}(p^*)}. \tag{13}$$

In the tables, we present results for $100 \times \Delta$SCORE, and therefore, for example, a value of 5 in the tables indicates that model M outperforms the benchmark by about 5% in density forecasting.

To provide a rough gauge of the statistical significance of differences in density scores, we use the Amisano and Giacomini (2007) t-test of equal means, applied to the log score for each model relative to the benchmark random forecast. We view the tests merely as a rough gauge because, with nested models, the asymptotic validity of the Amisano and Giacomini (2007) test requires that, as forecasting moves forward in time, the models be estimated with a rolling, rather than expanding, sample of data. As was the case for the point forecasts, we treat the tests as one-sided, and the t-statistics are computed with a serial correlation-robust variance, using a rectangular kernel, $h-1$ lags, and the small-sample adjustment of Harvey et al. (1997).

4. Results

We organize the discussion of our results around two main questions. First, we address the issue of the role of multivariate information in forecasting fiscal variables. We find that the use of a multivariate model does help, provided that shrinkage is imposed in order to reduce the parameter uncertainty. Second, we address the role of time variation in the coefficients and volatilities for forecasting. Our results indicate that time variation is particularly useful when producing density forecasts of fiscal variables.

4.1. The role of multivariate information

To assess the role of multivariate information in forecasting fiscal variables, we consider the BVARs estimated on the complete cross section of the nine variables under analysis and compare them with those of two alternative univariate models, the BAR benchmark and a RW forecast. The results of this comparison are shown in Tables 2 and 3. Specifically, Table 2 contains results for a point forecast evaluation and reports the RelRMSFE, as described by Eq. (12), while Table 3 contains results for a density forecast evaluation and reports the ΔSCORE, as described by Eq. (13). In both cases, these loss functions are computed against the benchmark BAR model described in Eq. (9).

Both tables are divided into four panels, reporting the results for the four countries under analysis. The best forecasting model for each variable-forecast horizon combination is highlighted in bold. The cases in which a model outperforms the BAR benchmark are highlighted in green. Finally, the symbols *, ** and *** denote rejection of the null of equal forecast accuracy at the 10%, 5% and 1% confidence levels.

We first focus on point forecasts, displayed in Table 2. Several conclusions can be drawn.

First, the large BVAR with constant coefficients outperforms the BAR benchmark in most instances. In particular, the BVAR produces the best point forecasts for all three fiscal variables when considering US data, with large and often significant gains. The gains against the BAR can go up to 18%, with the MAPE at the 4-quarter-ahead horizon being 2.24% for expenditures, 2.39% for revenues, and 6.48% for interest payments.

The BVAR outperforms the univariate benchmark for two fiscal variables out of three when considering data from Germany and France. The gains relative to the BAR can reach 28% for Germany and 12% for France. For France, the MAPEs produced by the BVAR at the 4-quarter-ahead horizon are 1.72% for expenditures, 1.37% for revenues, and 7.27% for interest payments. For Germany, the MAPEs produced by the BVAR at the 4-quarter-ahead horizon for the three fiscal variables are 3.52%, 2.11%, and 8.33%.

On the other hand, the results for the UK show that the BVAR outperforms the benchmark for expenditures, but is outperformed in forecasting the remaining two fiscal variables. The MAPEs produced by the BVAR at the 4-quarter-ahead horizon for the three fiscal variables are 2.50%, 2.60%, and 10.93%.

When comparing the performance with that of the simple RW model, the BVAR provides a better forecasting performance overall, outperforming the RW for all fiscal variables with US data, and for expenditure and revenues for the remaining countries. However, the RW does provide a competitive forecast of interest payments. It is worth noting that the RW no-change forecast can be thought of as the limit to which the BVAR converges as the shrinkage parameter approaches 0, meaning that using a tighter prior on the BVAR might improve the forecasts of interest payments.

Finally, looking also at the remaining variables in the system, the performance of the BVAR is generally superior to those of the BAR and RW in forecasting inflation and GDP growth, but is inferior when forecasting interest rates. The finding that a univariate specification seems to work better for interest rates is in line with several results in the literature, see e.g. Banbura et al. (2010), Carriero et al. (2013a), and Giannone et al. (2012).

We now turn our attention to density forecasts, displayed in Table 3.

As with point forecasts, the density forecasts provided by the BVAR with US data are the best overall, with large and often significant gains relative to the two alternative univariate specifications. The BVAR also produces the best density forecasts of expenditure for the UK, Germany and France, while it is generally outperformed by the BAR model in density forecasting for the remaining fiscal

Table 2
Point forecast evaluation of BVAR vs the benchmark.

horizon	Expenditures	Revenues	Int. payments	GDP growth	3-month rate	10-year rate	Inflation	Consumption	Industrial production
	US								
	RW vs benchmark								
1	1.02	1	0.96	0.91	**0.87** ***	**1**	1.09	**1.05**	1.17
2	1.05	0.98	0.95	0.92	**0.91** **	**0.98**	1.06	**1.09**	1.15
3	1.05	0.96	0.97	0.99	**0.98**	**1**	1.02	**1.13**	**1.11**
4	1.05	0.98	0.95	0.95	**0.98**	**1.01**	**0.89** *	1.17	**1.07**
	BVAR vs benchmark								
1	**0.92**	**0.89** *	**0.92** **	**0.82** **	1.16	1.07	**1.02**	1.07	**1.07**
2	**0.94**	**0.86** **	**0.87** ***	**0.93** **	1.08	1.08	**0.98**	1.14	**1.1**
3	**0.99**	**0.89** *	**0.84** ***	**0.94** **	1.06	1.07	**0.98**	1.15	**1.11**
4	**1.02**	**0.94**	**0.82** ***	**0.91** ***	1.05	1.08	1	**1.13**	1.11
	UK								
	RW vs benchmark								
1	1.02	**1.03**	**0.9**	1.02	**1.04**	**1.02**	1.13	1.03	**0.98**
2	1.06	1.06	**0.97** **	1.09	**1.04**	1.04	1.13	1.02	**0.97**
3	1.04	1.08	**0.92**	1.14	**1.04**	1.06	1.1	**1.03**	**0.96**
4	1.06	**1.04**	**0.96**	1.11	**1.06**	1.1	**1**	**1.03**	**0.97**
	BVAR vs benchmark								
1	**1**	**1.03**	0.99	**1.01**	1.18	1.1	**0.95**	**1**	1.1
2	**0.93**	**1**	1.02	**1.02**	1.1	**1.03**	**0.97**	**0.99**	1.08
3	**0.94** **	**1.04**	1.01	**1.04**	1.07	**1.03**	**1.06**	**1.03**	1.09
4	**0.94** *	1.06	1.03	**1.05**	**1.06**	**1.02**	1.08	1.06	1.08
	Germany								
	RW vs benchmark								
1	0.98	**0.61**	**1.04**	1.01	1.11	**0.95** *	1.27	1	**0.99**
2	0.97	**0.93**	**1.05**	1.25	1.09	**0.97**	1.37	0.99	**1**
3	**0.97**	1	**1.14**	1.25	**1.03**	**0.98**	1.16	1.02	**1.01**
4	**0.97**	0.79	**1**	1.39	1.03	**1**	1.52	1.06	**0.95**
	BVAR vs benchmark								
1	**0.89** *	0.72	1.2	**0.94**	**1.07**	1.06	**0.94**	**0.88**	1.02
2	**0.92** *	**0.93**	1.47	**0.97** *	**1.05**	1.04	**0.97**	**0.95**	1.05
3	**0.97**	**0.99**	1.39	**1**	1.04	1.04	**1**	**0.98**	1.11
4	1	**0.72**	1.55	**0.98**	**1.02**	1.05	**0.93** *	**1**	1.13
	France								
	RW vs benchmark								
1	1	1.16	**0.94**	0.99	1.03	**0.89** *	1.22	0.98	1.01
2	1.01	1.04	**0.89**	1.06	1.02	**0.89** *	1.14	1	0.99
3	0.99	**0.99**	**0.9**	1.27	0.98	**0.93**	1.19	1.01	0.96 *
4	0.97 *	1.13	**0.89**	1.28	0.98	**0.97**	1.31	0.99	0.97
	BVAR vs benchmark								
1	**0.94**	**1**	0.95 *	**0.94**	**0.89**	1.01	**0.9**	**0.96**	**0.85** *
2	**0.88**	**0.96**	0.9 *	**0.92**	**0.87**	1.07	**0.99**	**0.9**	**0.86** *
3	**0.88**	1.05	**0.9** *	**0.99**	**0.91** *	1.12	**0.98**	**0.89** *	**0.89** *
4	**0.89**	**1.04**	0.94	**1.02**	**0.92**	1.12	**0.96**	**0.9** *	**0.9**

Relative root mean squared forecast errors (RMSFEs) versus the BAR benchmark. A relative RMSFE value below one (highlighted in green) signals that a model outperforms the benchmark. Figures in bold denote the best model (within the VAR class) for each variable and forecast horizon. Gains in accuracy that are statistically different from zero are denoted by *, ** and ***, corresponding to significance levels of 10%, 5% and 1% respectively, evaluated using Diebold and Mariano (1995) *t*-statistics computed with a serial correlation-robust variance, using a rectangular kernel, $h - 1$ lags, and the small-sample adjustment of Harvey et al. (1997).

variables. As we shall see below, the BVAR performance for density forecasting improves dramatically following the inclusion of drifting parameters and volatilities.

Looking also at the remaining variables in the system, the BVAR generally provides the best density forecasts, outperforming both the BAR benchmark and the RW benchmark for most variables, and for inflation and GDP growth in particular (the latter with the exception of the UK). It is worth noting that the RW density forecasts are poor overall, due mainly to the fact that the RW

Table 3
Density forecast evaluation of BVAR vs the benchmark.

horizon	Expendi-tures	Revenues	Int. payments	GDP growth	3-month rate	10-year rate	Inflation	Consumption	Industrial production
	US								
	RW vs benchmark								
1	-13.1	-21.67	**20.49** *	-3.19	-42.77	-25.63	-4.76	-33.88	-17.43
2	-48.59	-54.88	**27.07**	-23.39	-89.77	-67.34	-11.77	-82.06	-23.14
3	-69.78	-76.21	**5.47**	-40.51	-123.72	-97.92	-16.7	-114	-33.56
4	-84.63	-87.07	-5.35	-48.81	-141.25	-121.6	-19.41	-138.07	-38.79
	BVAR vs benchmark								
1	**11.47** *	**11.54** *	10.23	**23.12** ***	**-2.32**	**-1.68**	**-1.9**	**3.99**	**-0.54**
2	**6.94**	**20.78** **	18.63 **	**10.92** ***	**-4.65**	**0.13**	**1.32**	**1.8**	**-3.76**
3	**3.14**	**16.16** *	**12.81** *	**2.67**	**-6.57**	**1.31**	**1.14**	**-0.61**	**-4.49**
4	**-2.69**	**0.17**	**10.58**	**8.01** ***	**-6.14**	**-0.91**	**0.81**	**-0.79**	**-5.75**
	UK								
	RW vs benchmark								
1	-40.49	-59.62	-3.09	-28.28	-28.98	-36.1	-21.23	-34.15	-17.35
2	-92.76	-118.95	-9.86	-51.7	-62.92	-78.48	-41.47	-81.79	-52.02
3	-120.11	-154.55	-40.97	-66.38	-88.72	-110.03	-54.78	-118.71	-72.91
4	-138.22	-180.4	-67.71	-70.3	-110.5	-133.85	-62.05	-143.51	-84.68
	BVAR vs benchmark								
1	**4.78** **	**-0.57**	**7.88**	**-0.66**	**-2.18**	**-1.09**	**7.87** ***	**3.26**	**-0.48**
2	**3.86** *	**0.71**	**-1.19**	**-4.65**	**-2.82**	**1.58**	**5.48** **	**2.96**	**-2.04**
3	**1.34**	**-2.69**	**-6.14**	**-5.66**	**-3.72**	**0.46**	**0.58**	**-1.82**	**-12.03**
4	**4.69** ***	**-5.44**	**-3.02**	**-1.91**	**-3.8**	**1.11**	**-0.5**	**-5.87**	**-9.73**
	Germany								
	RW vs benchmark								
1	-11.71	-48.3	-40.09	**12.42**	-23.36	**7.64**	-35.8	9.55	-28.79
2	-17.49	-126.87	-94.49	-22.29	-47.61	-23.29	-58.5	17.46	-76.83
3	-21.2	-162.52	-120.21	-19.75	-64.79	-53.57	-72.35	4.85	-114.95
4	-18.03	-191.8	-150.61	-37.65	-81.53	-76.12	-89.03	4.68	-143.34
	BVAR vs benchmark								
1	**6.33**	**0.19**	**-6.98**	5.84	**6.04**	-2.32	**6.62**	**12.72** *	**-2.18**
2	**7.95**	**-11.69**	**-10.31**	**4.03** **	**1.89**	**-1.78**	**1.94**	**18.37**	**-3.92**
3	**10.78**	**-9.5**	**-8.16**	**5.22**	**8.5**	**-2.46**	**0.81**	**9.24** *	**-9.89**
4	**10.78**	**-8.07**	**-12.3**	**3.84**	**4.25**	**-2.84**	**2.86**	**9.63**	**-11.65**
	France								
	RW vs benchmark								
1	2.24	-52.25	**8.52**	1.52	-27.58	**9.5**	-26.14	-0.02	-13.28
2	-8	-112.58	18.64	-12.9	-51.23	-18	-2.87	-16.92	-24.49
3	-20.53	-143.54	-27.04	-25.74	-81.03	-45.85	-18.34	-34.73	-34.98
4	-24.39	-171.75	-40.27	-36.07	-102.33	-69.31	-27.54	-40.92	-37.76
	BVAR vs benchmark								
1	**9.88** *	**-4.45**	-8.04	**4.58**	**14.56** ***	-2.7	**13.98**	**6.43**	**11.16**
2	**17.58**	**-3.53**	**29.67**	**4.02**	**24.1** **	**-4.57**	**15.29**	**14.37**	**16.36**
3	**22.52**	**-6.34**	**-2.02**	**1.2**	**11.49**	**-6.55**	**8.93**	**10.72**	**15.93**
4	**23.7**	**-8.38**	**9.27**	**-8.34**	**8.75** *	**-6.77**	**6.22**	**21.17**	**31.05**

Average difference in SCORE versus the BAR benchmark (multiplied by 100). A value above zero in the SCORE differences (highlighted in green) signals that a model outperforms the benchmark. As the SCOREs are measured in logs, a 100 x SCORE difference of 5, for example, signals a 5% gain in terms of density forecast accuracy. Figures in bold denote the best model (within the VAR class) for each variable and forecast horizon. Gains in accuracy that are statistically different from zero are denoted by *, ** and ***, corresponding to significance levels of 10%, 5% and 1% respectively, evaluated using Amisano and Giacomini (2007) *t*-statistics, computed with a serial correlation-robust variance, using a rectangular kernel, $h-1$ lags, and the small-sample adjustment of Harvey et al. (1997).

specification leads to predictive densities that are too spread out, especially at long horizons.

To summarize, our results show that the use of a large information set does help in forecasting fiscal variables. However, it should be noted that the mere use of a large information set is not sufficient, as some degree of shrinkage is also needed, to avoid the problem of overparameterization. To stress this point better, we have also considered repeating the forecasting exercise using a classical VAR. We obtained this model by imposing a

Table 4
Point forecast evaluation of VAR vs the benchmark.

horizon	Expenditures	Revenues	Int. payments	GDP growth	3-month rate	10-year rate	Inflation	Consumption	Industrial production
	US								
	VAR vs benchmark								
1	1.29	1.21	1.02	1.25	2.5	1.61	1.64	1.76	1.38
2	1.28	1.06	1	1.33	1.9	1.4	1.69	2.04	1.38
3	1.3	1.03	0.99	1.28	1.53	1.25	1.62	1.86	1.31
4	1.33	1.04	0.96	1.1	1.48	1.25	1.52	1.63	1.27
	UK								
	VAR vs benchmark								
1	1.42	1.61	1.06	1.39	1.9	1.87	1.42	1.4	1.39
2	1.25	1.36	1.03	1.23	1.59	1.7	1.16	1.33	1.28
3	1.29	1.38	1.03	1.48	1.45	1.55	1.4	1.31	1.24
4	1.39	1.42	1.04	1.46	1.39	1.43	1.43	1.31	1.22
	GE								
	VAR vs benchmark								
1	2.09	1.83	6.41	1.66	2.21	2.97	2.07	1.42	2.71
2	2.04	2.99	6.38	1.65	1.8	2.65	2.71	1.78	2.67
3	1.86	2.97	5.45	1.56	1.66	2.82	3.17	1.92	2.82
4	2.15	2.66	11.13	2.03	1.92	2.87	3.55	1.98	3.41
	FR								
	VAR vs benchmark								
1	1.43	2.98	1.02	1.99	2.38	2.18	2.05	1.14	2.88
2	1.15	3.57	0.97	1.71	1.33	3.11	1.81	1.1	1.86
3	1.26	4.1	0.98	1.98	1.3	3.97	2.45	1.19	1.74
4	1.15	4.63	1.59	1.73	1.25	4.55	2.59	1.1	1.64

Relative root mean squared forecast errors (RMSFEs) versus an unrestricted classical VAR. A relative RMSFE value above one (highlighted in green) signals that the VAR model is outperformed by the benchmark.

very loose prior on the BVAR, while the lag length was still chosen by maximizing the marginal likelihood.

The results for point forecasts[9] are displayed in Table 4, and show that the forecasts from a VAR with nine variables are systematically considerably inferior to forecasts based on a univariate specification, a result that is in line with the findings of Favero and Marcellino (2005). As our results from Tables 2 and 3 make clear, this is due, not to the fact that the larger data set does not contain any valuable information, but rather to the fact that the overparameterization inherent in the standard VAR framework makes it difficult to use such information.

4.2. The role of time variation in coefficients and volatilities

We now turn our attention to the role of time variation in the conditional mean coefficients and error volatilities. Tables 5a–8 provide the results of a forecasting exercise based on the TVP-BVAR described in Eq. (6) for the US, the UK, Germany, and France, respectively.

In the interest of space, we omit from these tables the results for the various alternative variables that are used as the fourth variable in the various TVP-BVAR specifications. However, the results for these variables broadly follow the same pattern as the fiscal variables, and are available upon request.

Each of the four tables is divided into two parts, (a) and (b), which show the results for the point and density forecasts respectively. Each of these parts is divided into four panels. Panel A contains the results based on the constant coefficients BVAR. These are the same figures as appear in Tables 2 and 3, and are included here to facilitate comparisons across models. Panel B contains the results obtained using the SV-BAR benchmark in Eq. (10), which is especially informative for density forecasting.

Panel C contains results from the TVP-BVARs. This panel contains the results of four-dimensional TVP-BVARs, where the three fiscal variables are always included in the model, and the fourth variable changes across specifications. Finally, Panel D contains the results from a version of the TVP-BVAR model where the variation in the conditional mean parameters Φ_t has been shut down, and only the volatility matrix $\Sigma_t^{1/2}$ changes over time. We label this model SV-BVAR, where SV stands for stochastic volatility.

[9] We did not compute density forecasts for the classical VAR because the overall parameter uncertainty in this model is so large that, in practice, a simulation of the whole predictive distribution is not feasible, due to the extremely high percentage of draws that fall in the nonstationary region, thus implying an explosive behaviour for the simulated paths of the variables. This difficulty per se warns against using large VARs without imposing some shrinkage on their parameters.

Table 5a
Point forecast evaluation, US data.

horizon	Expendi-tures	Revenues	Interest payments	Expendi-tures	Revenues	Interest payments
	PANEL A: BVAR vs benchmark			**PANEL B: SV-BAR vs benchmark**		
1	**0.92**	0.89 *	**0.92** **	1.03	1.00	1.06
2	**0.94**	0.86 **	**0.87** ***	1.03	1.01	1.10
3	0.99	0.89 *	**0.84** ***	1.02	1.01	1.09
4	1.02	0.94	**0.82** ***	1.02	1.01	1.06
	PANEL C: TVP-BVARs vs benchmatk			**PANEL D: SV-BVARs vs benchmatk**		
	4th-variable = GDP growth			4th-variable = GDP growth		
1	1.06	0.95	1.02	1.08	0.92 *	0.97
2	1.03	0.93 *	1.03	1.08	0.91 *	0.98
3	1.01	0.94 *	1.01	1.05	0.92 *	0.94 **
4	1.01	0.97	0.97	1.04	0.94	0.90 ***
	4th-variable=3-month rate			4th-variable=3-month rate		
1	1.02	**0.87** **	1.02	1.10	0.92 *	1.01
2	1.00	**0.82** **	1.02	1.07	0.91 *	1.02
3	0.99	**0.84** **	0.98	1.05	0.92 *	0.99
4	0.99	**0.86** *	0.94 ***	1.03	0.94	0.97 ***
	4th-variable=10-year rate			4th-variable=10-year rate		
1	1.07	0.93 *	1.09	1.10	0.92 *	1.02
2	1.03	0.92 *	1.11	1.07	0.91 *	1.02
3	1.01	0.93 *	1.08	1.05	0.92 *	1.00
4	1.00	0.95	1.04	1.03	0.94	0.97 ***
	4th-variable=inflation			4th-variable=inflation		
1	1.04	0.94	1.03	1.10	0.94	1.02
2	0.99	0.93 *	1.03	1.08	0.93 *	1.04
3	0.97	0.94	1.00	1.06	0.94	1.03
4	0.96	0.96	0.97 **	1.04	0.96	1.01
	4th-variable=consumption			4th-variable=consumption		
1	1.04	0.95	1.07	1.08	0.92 *	1.02
2	1.02	0.94 *	1.09	1.07	0.91 *	1.02
3	1.01	0.94	1.07	1.06	0.92 *	1.00
4	1.00	0.97	1.04	1.05	0.94	0.97 ***
	4th-variable=industrial production			4th-variable=industrial production		
1	0.99	0.91 *	1.07	1.01	0.91 *	1.01
2	0.95	0.88 **	1.08	0.99	0.88 **	1.01
3	**0.94**	0.88 **	1.08	0.98	0.88 **	0.99
4	**0.94**	0.91	1.06	0.97	0.91	0.97 *

Relative root mean squared forecast errors (RMSFEs) versus the BAR benchmark. A relative RMSFE value below one (highlighted in green) signals that a model outperforms the benchmark. Figures in bold denote the best model (within the VAR class) for each variable and forecast horizon. Gains in accuracy that are statistically different from zero are denoted by *, ** and ***, corresponding to significance levels of 10%, 5% and 1% respectively, evaluated using Diebold and Mariano (1995) t-statistics computed with a serial correlation-robust variance, using a rectangular kernel, $h-1$ lags, and the small-sample adjustment of Harvey et al. (1997).

In the tables, a green shade on a given variable-forecast horizon combination signals that the model improves on the BAR benchmark. The best model for each variable-forecast horizon combination is highlighted in bold.

4.2.1. Results for the US

Tables 5a and 5b show the results for the US, for point and density forecasts respectively. Focusing first on point forecasts, the TVP-BVAR models provide good point

Table 5b
Density forecast evaluation, US data.

horizon	Expendi-tures	Revenues	Interest payments	Expendi-tures	Revenues	Interest payments
	PANEL A: BVAR vs benchmark			**PANEL B: SV-BAR vs benchmark**		
1	11.47 *	11.54 *	10.23	4.18	**31.89** ***	64.85 ***
2	6.94	20.78 **	18.63 **	2.51	23.97 ***	46.44 **
3	3.14	16.16 *	12.81 *	2.79	21.64 ***	16.74
4	-2.69	0.17	10.58	3.01	25.80 ***	6.15
	PANEL C: TVP-BVARs vs benchmatk			**PANEL D: SV-BVARs vs benchmatk**		
	4th-variable = GDP growth			4th-variable = GDP growth		
1	2.46	20.39 **	67.10 ***	-4.65	10.27	69.17
2	2.26	23.78 **	57.50 **	-5.17	21.14 **	57.92 **
3	4.46	22.48 ***	30.21 *	-2.12	15.80 **	**37.06** **
4	3.99	23.59 **	21.89 **	-3.72	25.37 **	**29.26** **
	4th-variable=3-month rate			4th-variable=3-month rate		
1	0.83	24.36 ***	72.69 ***	0.35	20.50	66.34
2	0.33	**36.62** ***	58.00 **	-2.46	22.23 **	50.14 **
3	1.11	24.17 **	28.87 *	-0.55	20.53 ***	28.67 *
4	-0.56	**37.58** ***	17.91 *	-5.83	27.12 ***	22.44
	4th-variable=10-year rate			4th-variable=10-year rate		
1	3.56	24.54	**74.39**	-4.01	21.46	65.79
2	-1.38	25.57 ***	**60.43** ***	-2.46	29.47 **	50.24 **
3	-0.47	21.60 ***	33.24 **	-1.89	20.65 ***	28.39 *
4	0.33	25.73	19.00	-6.08	30.08	22.35
	4th-variable=inflation			4th-variable=inflation		
1	-5.00	19.87 **	63.84 ***	-4.76	24.84	71.43
2	1.03	26.71 ***	53.12 **	-1.78	21.99 **	46.57 **
3	5.87	16.27 **	30.03 **	-0.07	16.80 **	25.67
4	5.51	23.11	24.37	-3.14	35.23	19.14
	4th-variable=consumption			4th-variable=consumption		
1	-1.72	17.99 **	65.74 ***	-2.70	15.50	66.45
2	0.87	18.96 **	50.93 **	-0.01	22.15 **	49.32 **
3	2.08	**29.90** ***	24.54	-0.64	20.54 ***	27.79 *
4	3.90	23.16	17.23	-0.91	26.99	23.08
	4th-variable=industrial production			4th-variable=industrial production		
1	**16.56** *	27.29 ***	71.23 ***	7.07	24.07 *	69.70 ***
2	**8.55**	36.44 ***	55.17 **	4.09	32.90 **	53.78 **
3	**8.62**	25.12 ***	26.94 *	2.46	23.10 ***	31.02 *
4	**6.35**	26.28 ***	15.99	-0.66	31.99	22.44

Average difference in SCORE versus the BAR benchmark (multiplied by 100). A value above zero in the SCORE differences (highlighted in green) signals that a model outperforms the benchmark. As the SCOREs are measured in logs, a 100 x SCORE difference of 5, for example, signals a 5% gain in terms of density forecast accuracy. Figures in bold denote the best model (within the VAR class) for each variable and forecast horizon. Gains in accuracy that are statistically different from zero are denoted by *, ** and ***, corresponding to significance levels of 10%, 5% and 1% respectively, evaluated using Amisano and Giacomini (2007) *t*-statistics computed with a serial correlation-robust variance, using a rectangular kernel, $h-1$ lags, and the small-sample adjustment of Harvey et al. (1997).

forecasts for revenues, systematically outperforming the BAR benchmark, while only some specifications provide gains for the other two fiscal variables. For expenditures, the model with industrial production outperforms the BAR, while the TVP-BVAR is outperformed by the benchmark overall for interest payments. On the other hand, BVAR based on the larger information set systematically produces good forecasts for all three fiscal variables, outperforming the benchmark in all but one case (forecasts of expenditures four quarters ahead). The BVAR produces the best point forecasts overall, with the exception of revenues, which are predicted slightly better by the the TVP-BVAR with the 3-month rate as the fourth variable. These results indicate that, for point forecasts, the use of a large information set is of more importance than the modelling of time variation in the coefficients and volatilities, and that the use of a large BVAR should be preferred to the use of smaller systems with time-varying coefficients.

Turning to density forecasts, both the BVAR and the TVP-BVARs improve on the BAR benchmark in most cases. The TVP-BVARs outperform the BVAR in density forecasting, producing systematically higher scores for revenues and interest payments, and often higher scores for expenditures. Both models improve on the univariate SV-BAR benchmark, showing that the joint modelling of macroeconomic variables improves the density forecasts.

By comparing the results in Panel C with those in Panel D, we can assess the relative contributions of time variation in the coefficients and in the volatilities to the improvement in density forecasts. The TVP-BVARs provide better forecasts than the SV-BVARs in 23 cases out of 24 for expenditures, in 15 cases out of 24 for revenues, and in 12 cases out of 24 for interest payments. Therefore, the use of time variation in both the coefficients and the volatilities does help in the density forecasting of fiscal variables, more so than does variation in the volatilities alone. Note this is not necessarily the case for point forecasts, where the results are mixed, with the TVP-BVARs providing better results for expenditures, the SV-BVARs being better for interest payments, and the two models providing similar forecasts for revenues.

To summarise, the results for the US show that the use of a large information set improves the point forecasts, while the use of time variation in both coefficients and volatilities can improve the density forecasts.

4.2.2. Results for the UK

Tables 6a and 6b show results for the UK, for point and density forecasts respectively. Focusing first on point forecasts, the TVP-BVARs do not provide good point forecasts of the fiscal variables, and none of the TVP-BVAR specifications are able to outperform either the benchmarks or the large BVAR. This confirms the finding that the use of a large information set helps relatively more in point forecasting than the use of a model with variation in the coefficients. However, the evidence in favour of a multivariate model over a univariate model is somewhat less strong for the UK than for the US. The best point forecasts for the UK are obtained by the BVAR (for expenditures), the univariate SV-BAR (for interest payments), and the univariate BAR benchmark (for revenues).

Turning to density forecasts, as with the US data, both the BVAR and the TVP-BVARs improve on the BAR benchmark in most cases. In particular, the TVP-BVARs produce the best density forecasts of interest payments overall. Most of the TVP-BVAR specifications also produce good forecasts of revenues, but even better forecasts for this variable are produced by the SV-BAR. The poor performance of the TVP-BVARs for the density forecasting of expenditures at long horizons is driven by the poor performances of these models in producing good point forecasts for this variable.

By comparing the results in Panel C with those in Panel D, we can assess the relative contributions of time variation in the coefficients and in the volatilities to the improvement in density forecasts. The TVP-BVARs provide better forecasts than the SV-BVARs in 20 cases out of 24 for expenditures, in all 24 cases for revenues, and in 17 cases out of 24 for interest payments. Therefore, the usefulness of time variation in both the conditional mean coefficients and the volatilities (rather than in the volatilities alone) is even stronger using UK data.

To summarise, the results for the UK provide weaker evidence in favour of the use of multivariate models, but the result that, among the multivariate specifications, the large BVAR outperforms the TVP-BVARs in point forecasting still holds. Moreover, as was the case with the US data, we find that modelling the time variation in both the coefficients and the volatilities improves the density forecasts.

4.2.3. Results for Germany and France

We now turn to the results for Germany and France. It is important to stress that caution is needed when interpreting the results for these countries, as the data-set available includes only about 20 years of quarterly data (from 1991 to the end of 2010). This caveat is particularly important for the time-varying specifications, as the sample is further reduced in this case, because the first five years of data are used as a training sample.

Tables 7a and 7b show the results for Germany, for point and density forecasts respectively. Focusing first on the point forecasts, the BVAR based on the larger information set systematically produces good forecasts, beating the BAR benchmark for two out of the three fiscal variables, expenditures and revenues. No model is able to beat the BAR benchmark for interest payments. The BVAR is the best model overall in forecasting expenditures, but the univariate SV-BAR produces the best forecasts for revenues. The TVP-BVARs, on the other hand, are never able to outperform the benchmark for point forecasting. This latter result is probably due to the limited size of the sample used, which is too small to provide good estimates of time-varying specifications.

Turning to density forecasts, the TVP-BVAR models perform well overall for forecasting revenues and interest payments, as does the BVAR for forecasting expenditures. By comparing the results contained in Panel C with those in Panel D, it is clear that the use of time variation in the conditional mean coefficients helps. In particular, the TVP-BVARs provide better forecasts than the SV-BVARs in 20 cases out of 24 for expenditures, in all 24 cases for revenues, and in 17 cases out of 24 for interest payments. Therefore, as with the US and UK data, the use of time variation in both the conditional mean coefficients and the volatilities does improve the density forecasts of fiscal variables, relative to variation in the volatilities alone.

Tables 8a and 8b show the results for France, for point and density forecasts respectively. Focusing first on point forecasts, the BVAR based on the larger information set systematically produces good forecasts for all three fiscal

Table 6a
Point forecast evaluation, UK data.

horizon	Expendi-tures	Revenues	Interest payments	Expendi-tures	Revenues	Interest payments
	PANEL A: BVAR vs benchmark			PANEL B: SV-BAR vs benchmark		
1	**1.00**	1.03	0.99	1.03	**1.00**	**0.96**
2	**0.93**	**1.00**	1.02	1.05	**1.00**	**0.99**
3	**0.94** **	1.04	1.01	1.06	**1.00**	**0.99**
4	**0.94** *	1.06	1.03	1.07	**1.00**	**0.99**
	PANEL C: TVP-BVARs vs benchmatk			PANEL D: SV-BVARs vs benchmatk		
	4th-variable = GDP growth			4th-variable = GDP growth		
1	1.04	1.06	0.97	1.04	1.05	1.01
2	1.03	1.04	1.00	1.06	1.04	1.03
3	1.03	1.06	1.01	1.08	1.06	1.06
4	1.03	1.07	1.03	1.10	1.08	1.08
	4th-variable=3-month rate			4th-variable=3-month rate		
1	1.09	1.04	0.97	1.09	1.02	0.98
2	1.09	1.03	0.99	1.12	1.02	1.01
3	1.11	1.04	1.00	1.15	1.03	1.02
4	1.12	1.04	1.02	1.16	1.03	1.04
	4th-variable=10-year rate			4th-variable=10-year rate		
1	1.07	1.04	0.97	1.09	1.02	0.97
2	1.09	1.03	1.00	1.11	1.02	1.01
3	1.11	1.04	1.01	1.15	1.03	1.02
4	1.12	1.04	1.02	1.16	1.03	1.04
	4th-variable=inflation			4th-variable=inflation		
1	1.09	1.06	0.96	1.08	1.03	0.99
2	1.09	1.05	1.00	1.12	1.03	1.02
3	1.11	1.06	1.00	1.15	1.04	1.03
4	1.11	1.07	1.02	1.16	1.05	1.05
	4th-variable=consumption			4th-variable=consumption		
1	1.10	1.05	0.98	1.09	1.02	0.99
2	1.10	1.03	1.00	1.11	1.01	1.02
3	1.11	1.04	1.01	1.14	1.02	1.03
4	1.11	1.04	1.03	1.15	1.03	1.05
	4th-variable=industrial production			4th-variable=industrial production		
1	1.01	1.08	0.99	1.00	1.05	0.99
2	1.01	1.08	1.01	1.02	1.05	1.00
3	1.03	1.08	1.01	1.05	1.06	1.01
4	1.04	1.08	1.03	1.07	1.06	1.02

Relative root mean squared forecast errors (RMSFEs) versus the BAR benchmark. A relative RMSFE value below one (highlighted in green) signals that a model outperforms the benchmark. Figures in bold denote the best model (within the VAR class) for each variable and forecast horizon. Gains in accuracy that are statistically different from zero are denoted by *, ** and ***, corresponding to significance levels of 10%, 5% and 1% respectively, evaluated using Diebold and Mariano (1995) *t*-statistics computed with a serial correlation-robust variance, using a rectangular kernel, $h-1$ lags, and the small-sample adjustment of Harvey et al. (1997).

variables, outperforming the benchmark and the univariate BAR-SV in all but two cases, namely the forecasts of revenues at three and four quarters ahead. On the other hand, the TVP-BVAR models are never able to outperform the benchmark. In general, the models with time variation do not perform well for point forecasting: also, among the univariate specifications, the SV-BAR is outperformed systematically by the constant coefficient BAR. As was the case with the German data, this result is probably due to the limited size of the sample used.

Table 6b
Density forecast evaluation, UK data.

horizon	Expenditures	Revenues	Interest payments	Expenditures	Revenues	Interest payments
	PANEL A: BVAR vs benchmark			**PANEL B: SV-BAR vs benchmark**		
1	4.78 **	-0.57	7.88	4.40	**4.47** *	53.83 *
2	3.86 *	0.71	-1.19	-0.95	3.18	82.33
3	**1.34**	-2.69	-6.14	-5.10	**5.83** **	53.09 *
4	**4.69** ***	-5.44	-3.02	-7.61	**3.83**	50.25 *
	PANEL C: TVP-BVARs vs benchmatk			**PANEL D: SV-BVARs vs benchmatk**		
	4th-variable = GDP growth			4th-variable = GDP growth		
1	**6.31**	-0.74	48.57	3.51	-7.34	44.45
2	1.13	0.77	**91.06**	-6.73	-6.74	68.44
3	-2.27	0.73	**57.55**	-9.72	-6.64	50.01
4	-14.87	-3.32	**53.87**	-15.63	-11.23	44.27
	4th-variable=3-month rate			4th-variable=3-month rate		
1	1.75	1.18	46.52	2.38	-3.30	45.79
2	-3.57	3.05	82.36	-6.80	-1.19	72.53
3	-6.55	4.62	45.96	-10.09	-1.27	51.12
4	-19.62	1.33	37.00	-15.26	-1.46	40.43
	4th-variable=10-year rate			4th-variable=10-year rate		
1	2.35	-0.85	40.28	-1.09	-5.65	44.49
2	-1.38	**3.36**	70.57	-7.51	-1.62	70.20
3	-8.85	4.57	54.14	-11.03	-3.58	49.87
4	-12.42	2.18	38.35	-18.18	-2.95	38.77
	4th-variable=inflation			4th-variable=inflation		
1	1.37	-3.54	**53.95**	-1.76	-6.12	43.48
2	-3.72	1.53	90.16	-11.97	-3.48	66.34
3	-8.81	-1.59	53.77	-15.56	-5.53	48.03
4	-13.03	-4.54	43.11	-20.44	-6.79	34.72
	4th-variable=consumption			4th-variable=consumption		
1	1.52	-1.21	44.88	0.31	-4.12	41.34
2	-0.97	2.02	82.67	-5.43	-1.60	71.42
3	-25.18	1.42	47.48	-10.50	-3.18	48.71
4	-45.56	0.31	44.17	-15.43	-4.29	38.67
	4th-variable=industrial production			4th-variable=industrial production		
1	6.07	-2.76	45.19	4.84	-7.21	46.70
2	**3.95**	-2.94	83.95	-2.49	-3.80	70.04
3	-3.80	2.47	51.46	-8.03	-5.12	52.37
4	-6.71	1.93	49.92	-11.66	-5.58	40.21

Average difference in SCORE versus the BAR benchmark (multiplied by 100). A value above zero in the SCORE differences (highlighted in green) signals that a model outperforms the benchmark. As the SCOREs are measured in logs, a 100 x SCORE difference of 5, for example, signals a 5% gain in terms of density forecast accuracy. Figures in bold denote the best model (within the VAR class) for each variable and forecast horizon. Gains in accuracy that are statistically different from zero are denoted by *, ** and ***, corresponding to significance levels of 10%, 5% and 1% respectively, evaluated using Amisano and Giacomini (2007) *t*-statistics computed with a serial correlation-robust variance, using a rectangular kernel, $h-1$ lags, and the small-sample adjustment of Harvey et al. (1997).

Turning to density forecasts, there is no clear ranking across models. Either the BVAR or the TVP-BVARs improve over the BAR benchmark for expenditures and interest payments, while for revenues no model can beat the univariate BAR benchmark. By comparing the results contained in Panel C with those in Panel D, it is clear that the use of time variation in the conditional mean coefficients helps. In particular, the TVP-BVARs provide better forecasts than the SV-BVARs in 8 cases out of 24 for expenditures, in 20 cases out of 24 for revenues, and in 22

Table 7a
Point forecast evaluation, Germany.

horizon	Expendi-tures	Revenues	Interest payments	Expendi-tures	Revenues	Interest payments
	PANEL A: BVAR vs benchmark			PANEL B: SV-BAR vs benchmark		
1	**0.89** *	0.72	1.20	0.98	**0.60**	1.04
2	**0.92** *	0.93	1.47	0.97 *	**0.72**	1.12
3	**0.97**	0.99	1.39	0.98	**0.77**	1.16
4	1.00	**0.72**	1.55	**0.99**	0.76	1.23
	PANEL C: TVP-BVARs vs benchmatk			PANEL D: SV-BVARs vs benchmatk		
	4th-variable = GDP growth			4th-variable = GDP growth		
1	1.00	1.10	1.25	1.04	1.05	1.01
2	0.96	1.07	1.28	1.06	1.04	1.03
3	0.97	1.08	1.33	1.08	1.06	1.06
4	**0.99**	1.10	1.35	1.10	1.08	1.08
	4th-variable=3-month rate			4th-variable=3-month rate		
1	1.08	1.03	1.03	1.09	1.02	0.98
2	1.03	1.01	1.01	1.12	1.02	1.01
3	1.03	1.02	1.00	1.15	1.03	1.02
4	1.02	1.02	**0.99**	1.16	1.03	1.04
	4th-variable=10-year rate			4th-variable=10-year rate		
1	1.12	1.07	1.04	1.09	1.02	**0.97**
2	1.05	1.04	**0.99**	1.11	1.02	1.01
3	1.05	1.05	**1.00**	1.15	1.03	1.02
4	1.04	1.05	1.01	1.16	1.03	1.04
	4th-variable=inflation			4th-variable=inflation		
1	1.16	1.12	1.06	1.08	1.03	0.99
2	1.10	1.10	1.04	1.12	1.03	1.02
3	1.08	1.10	1.07	1.15	1.04	1.03
4	1.07	1.15	1.07	1.16	1.05	1.05
	4th-variable=consumption			4th-variable=consumption		
1	1.20	1.11	1.03	1.09	1.02	0.99
2	1.08	1.07	1.03	1.11	1.01	1.02
3	1.06	1.05	1.02	1.14	1.02	1.03
4	1.04	1.06	1.03	1.15	1.03	1.05
	4th-variable=industrial production			4th-variable=industrial production		
1	1.10	1.08	1.13	1.00	1.05	0.99
2	1.06	1.09	1.13	1.02	1.05	1.00
3	1.05	1.10	1.19	1.05	1.06	1.01
4	1.07	1.16	1.19	1.07	1.06	1.02

Relative root mean squared forecast errors (RMSFEs) versus the BAR benchmark. A relative RMSFE value below one (highlighted in green) signals that a model is outperforming the benchmark. Figures in bold denote the best model (within the VAR class) for each variable and forecast horizon. Gains in accuracy that are statistically different from zero are denoted by *, ** and ***, corresponding to significance levels of 10%, 5% and 1% respectively, evaluated using Diebold and Mariano (1995) *t*-statistics computed with a serial correlation-robust variance, using a rectangular kernel, $h-1$ lags, and the small-sample adjustment of Harvey et al. (1997).

cases out of 24 for interest payments. This confirms the results found using the data for the US, the UK, and Germany, namely that the use of time variation in both the coefficients and the volatilities does help in the density forecasting of fiscal variables, more so than the use of time varying volatilities only.

To summarise, the results for Germany and France broadly confirm the pattern found using the US and UK

Table 7b
Density forecast evaluation, Germany.

horizon	Expendi-tures	Revenues	Interest payments	Expendi-tures	Revenues	Interest payments
	PANEL A: BVAR vs benchmark			**PANEL B: SV-BAR vs benchmark**		
1	**6.33**	0.19	-6.98	-20.74	**9.49**	10.75
2	**7.95**	-11.69	-10.31	-4.75	-11.76	17.45 **
3	**10.78**	-9.50	-8.16	0.37	-4.64	9.25
4	10.78	-8.07	-12.30	**12.43**	-5.81	1.68
	PANEL C: TVP-BVARs vs benchmatk			**PANEL D: SV-BVARs vs benchmatk**		
	4th-variable = GDP growth			4th-variable = GDP growth		
1	6.31	-0.74	48.57	3.51	-7.34	44.45
2	1.13	0.77	**91.06**	-6.73	-6.74	68.44
3	-2.27	0.73	**57.55**	-9.72	-6.64	50.01
4	-14.87	-3.32	**53.87**	-15.63	-11.23	44.27
	4th-variable=3	0.00	0.00	4th-variable=3-month rate		
1	1.75	1.18	46.52	2.38	-3.30	45.79
2	-3.57	3.05	82.36	-6.80	-1.19	72.53
3	-6.55	4.62	45.96	-10.09	-1.27	51.12
4	-19.62	1.33	37.00	-15.26	-1.46	40.43
	4th-variable=1	0	0	4th-variable=10-year rate		
1	2.35	-0.85	40.28	-1.09	-5.65	44.49
2	-1.38	**3.36**	70.57	-7.51	-1.62	70.20
3	-8.85	**4.57**	54.14	-11.03	-3.58	49.87
4	-12.42	**2.18**	38.35	-18.18	-2.95	38.77
	4th-variable=in	0.00	0.00	4th-variable=inflation		
1	1.37	-3.54	**53.95**	-1.76	-6.12	43.48
2	-3.72	1.53	90.16	-11.97	-3.48	66.34
3	-8.81	-1.59	53.77	-15.56	-5.53	48.03
4	-13.03	-4.54	43.11	-20.44	-6.79	34.72
	4th-variable=c	0	0	4th-variable=consumption		
1	1.52	-1.21	44.88	0.31	-4.12	41.34
2	-0.97	2.02	82.67	-5.43	-1.60	71.42
3	-25.18	1.42	47.48	-10.50	-3.18	48.71
4	-45.56	0.31	44.17	-15.43	-4.29	38.67
	4th-variable=in	0.00	0.00	4th-variable=industrial production		
1	6.07	-2.76	45.19	4.84	-7.21	46.70
2	3.95	-2.94	83.95	-2.49	-3.80	70.04
3	-3.80	2.47	51.46	-8.03	-5.12	52.37
4	-6.71	1.93	49.92	-11.66	-5.58	40.21

Average difference in SCORE versus the BAR benchmark (multiplied by 100). A value above zero in the SCORE differences (highlighted in green) signals that a model outperforms the benchmark. As the SCOREs are measured in logs, a 100 x SCORE difference of 5, for example, signals a 5% gain in terms of density forecast accuracy. Figures in bold denote the best model (within the VAR class) for each variable and forecast horizon. Gains in accuracy that are statistically different from zero are denoted by *, ** and ***, corresponding to significance levels of 10%, 5% and 1% respectively, evaluated using Amisano and Giacomini (2007) *t*-statistics computed with a serial correlation-robust variance, using a rectangular kernel, $h-1$ lags, and the small-sample adjustment of Harvey et al. (1997).

data. The use of a large information set improves the point forecasts, while the use of time variation in both coefficients and volatilities improves the density forecasts. However, we did find that the time-varying specifications have worse point forecasting performances for France and Germany than for the US and the UK, probably due to the short length of the estimation sample. We therefore suggest the use of a constant coefficients specification whenever sufficiently long time series are not available.

Table 8a
Point forecast evaluation, France.

horizon	Expendi-tures	Revenues	Interest payments	Expendi-tures	Revenues	Interest payments
	PANEL A: BVAR vs benchmark			**PANEL B: SV-BAR vs benchmark**		
1	**0.94**	**1.00**	**0.95** *	1.03	**1.00**	1.30
2	**0.88**	**0.96**	**0.90** *	1.02	0.98	1.39
3	**0.88**	1.05	**0.90** *	1.02	**0.98**	1.49
4	**0.89**	1.04	**0.94**	1.02	**0.98**	1.60
	PANEL C: TVP-BVARs vs benchmatk			**PANEL D: SV-BVARs vs benchmatk**		
	4th-variable = GDP growth			4th-variable = GDP growth		
1	1.10	1.06	1.18	0.97	1.14	1.31
2	1.05	1.06	1.18	0.93	1.15	1.40
3	1.05	1.08	1.24	0.95	1.15	1.50
4	1.07	1.10	1.26	0.96	1.20	1.61
	4th-variable=3-month rate			4th-variable=3-month rate		
1	1.09	1.01	1.04	1.04	1.06	1.30
2	1.05	1.01	1.03	1.01	1.08	1.39
3	1.05	1.00	1.04	1.01	1.13	1.48
4	1.05	1.02	1.03	1.01	1.15	1.59
	4th-variable=10-year rate			4th-variable=10-year rate		
1	1.13	1.08	1.06	1.03	1.09	1.30
2	1.07	1.06	1.03	1.01	1.12	1.38
3	1.06	1.06	1.04	1.01	1.16	1.48
4	1.06	1.07	1.03	1.02	1.18	1.59
	4th-variable=inflation			4th-variable=inflation		
1	1.19	1.20	1.05	1.05	1.15	1.30
2	1.12	1.20	1.07	1.02	1.16	1.43
3	1.10	1.22	1.11	1.01	1.15	1.62
4	1.09	1.31	1.12	1.02	1.22	1.82
	4th-variable=consumption			4th-variable=consumption		
1	1.18	1.14	0.97	1.08	1.09	1.26
2	1.07	1.11	0.97	1.06	1.16	1.39
3	1.05	1.10	0.97	1.08	1.24	1.63
4	1.04	1.13	0.98	1.11	1.37	2.07
	4th-variable=industrial production			4th-variable=industrial production		
1	1.12	1.12	1.17	1.01	1.11	1.44
2	1.04	1.06	1.17	0.97	1.14	1.89
3	1.07	1.07	1.20	0.97	1.13	2.87
4	1.08	1.07	1.21	0.97	1.20	4.79

Relative root mean squared forecast errors (RMSFEs) versus the BAR benchmark. A relative RMSFE value below one (highlighted in green) signals that a model outperforms the benchmark. Figures in bold denote the best model (within the VAR class) for each variable and forecast horizon. Gains in accuracy that are statistically different from zero are denoted by *, ** and ***, corresponding to significance levels of 10%, 5% and 1% respectively, evaluated using Diebold and Mariano (1995) *t*-statistics computed with a serial correlation-robust variance, using a rectangular kernel, $h-1$ lags, and the small-sample adjustment of Harvey et al. (1997).

4.3. Results during recessions

Given the recent economic turmoil in the countries under consideration, it is interesting to consider how our forecasting models perform during recessions. We therefore re-calculate the forecast comparison statistics over the recession periods that occur during the forecast sample. The recession dates for the US are obtained from

Table 8b
Density forecast evaluation, France.

horizon	Expendi-tures	Revenues	Interest payments	Expendi-tures	Revenues	Interest payments
	PANEL A: BVAR vs benchmark			**PANEL B: SV-BAR vs benchmark**		
1	**9.88**	**-4.45**	-8.04	-2.03	-11.91	7.01
2	17.58	**-3.53**	29.67	**18.83**	-12.77	5.80
3	22.52	**-6.34**	-2.02	21.10	-12.23	-45.59
4	23.70	**-8.38**	9.27	**41.93**	-15.79	-49.70
	PANEL C: TVP-BVARs vs benchmatk			**PANEL D: SV-BVARs vs benchmatk**		
	4th-variable = GDP growth			4th-variable = GDP growth		
1	1.71	-21.14	24.59	1.88	-36.17	27.64
2	7.01	-26.08	31.67	1.10	-46.99	-3.42
3	4.57	-31.53	-17.61	23.96	-49.64	-139.49
4	0.80	-35.89	-16.78	4.55	-77.04	-90.61
	4th-variable=3-month rate			4th-variable=3-month rate		
1	-5.81	-17.65	**42.60**	-8.91	-67.90	27.75
2	3.91	-17.66	36.78	-1.45	-110.75	-8.14
3	0.31	-19.18	-3.80	-11.53	-129.77	-44.03
4	-4.29	-24.94	3.11	-13.78	-149.35	-52.24
	4th-variable=10-year rate			4th-variable=10-year rate		
1	-8.87	-15.50	31.29	4.77	-31.21	-7.90
2	4.28	-15.40	37.58	12.28	-40.24	-4.22
3	1.66	-22.29	-1.50	**29.49**	-43.25	-44.97
4	5.75	-24.81	4.61	23.65	-52.14	-57.50
	4th-variable=inflation			4th-variable=inflation		
1	-19.35	-24.25	39.82	-4.47	-35.92	28.36
2	-3.68	-22.61	44.99	6.03	-46.83	7.06
3	-4.53	-30.59	-6.27	5.96	-53.79	-52.01
4	17.76	-34.14	-2.51	14.77	-64.61	-63.10
	4th-variable=consumption			4th-variable=consumption		
1	-14.41	-20.68	42.89	-1.02	-32.18	24.28
2	4.59	-18.92	**46.09**	-3.40	-42.28	-6.74
3	-0.65	-22.67	**14.39**	-5.86	-46.43	-64.14
4	-5.65	-24.75	**15.48**	1.39	-56.70	-84.04
	4th-variable=industrial production			4th-variable=industrial production		
1	-46.93	-57.05	18.90	7.83	-34.58	31.08
2	-51.74	-71.51	12.36	15.32	-42.85	-4.84
3	-54.32	-77.40	-31.32	13.35	-44.23	-61.20
4	-46.86	-83.95	-29.32	25.86	-54.67	-85.28

Average difference in SCORE versus the BAR benchmark (multiplied by 100). A value above zero in the SCORE differences (highlighted in green) signals that a model outperforms the benchmark. As the SCOREs are measured in logs, a 100 x SCORE difference of 5, for example, signals a 5% gain in terms of density forecast accuracy. Figures in bold denote the best model (within the VAR class) for each variable and forecast horizon. Gains in accuracy that are statistically different from zero are denoted by *, ** and ***, corresponding to significance levels of 10%, 5% and 1% respectively, evaluated using Amisano and Giacomini (2007) *t*-statistics computed with a serial correlation-robust variance, using a rectangular kernel, $h-1$ lags, and the small-sample adjustment of Harvey et al. (1997).

the NBER, while OECD recession estimates are used for the remaining countries. For the sake of brevity, we merely summarise the key results in this section, and will make the detailed tables available on request.

4.3.1. BVARs

The performances of the point forecasts of fiscal variables from BVARs over recession periods are similar to the full sample results for all countries. However, there is some evidence that the density forecasts for some variables are more accurate during recessions. For example, for the US and France, the relative score for government expenditures and revenues calculated over recessions is larger than the full sample estimates. Similar results can be observed for the density forecasts of expenditures and interest payments for the UK and expenditures for Germany.

4.3.2. TVP-BVARs and SV-BVARs

As in the case of the constant coefficient VARs, there is little evidence to suggest that the point forecasts from TVP-BVARs and SV-BVARs improve substantially over recession periods, with the estimated relative root mean squared errors being similar to those obtained using the full forecast period. In contrast, the improvement in the accuracy of density forecasts from these models over the benchmark model displays a substantial increase during recession periods. The magnitude of this increase is generally much larger than that observed for the fixed coefficient BVAR. For the US, this improvement occurs for all fiscal variables, and is largest for interest payments. Similarly, for the UK and France, the estimated ΔSCORE for interest payments is much larger during recession periods. Finally, for Germany, the ΔSCORE for expenditures is larger during recession periods. Note that these improvements in ΔSCORE are of similar magnitudes across the TVP-BVARs and the SV-BVARs. This suggests that the stochastic volatility plays a key role during recessions, possibly reflecting the impact of heightened economic uncertainty during these periods.

4.4. Discussion

We have provided a large set of results for both point and density forecasts, and while we have found several differences across countries and variables, four main messages can be taken home from our forecasting exercise. First, we found that the use of multivariate models (VARs) helps in forecasting, in contrast to the conclusions of Favero and Marcellino (2005). This is due to the fact that using Bayesian shrinkage allows the problem of overparameterization to be reduced drastically. Second, we found that the use of a large information set generally helps, and rarely harms, the point forecasting of fiscal variables. This is in line with several contributions in the literature that have pointed towards the use of larger systems for point forecasting (Banbura et al., 2010; Carriero et al., 2013a; Koop, 2013). Third, we found that the use of drifting coefficients and volatilities does improve the density forecasts of fiscal variables, even though these models generally produce worse point forecasts than larger specifications with constant coefficients. This suggests that a large model with time variation in both the volatilities and coefficients would be the best, but unfortunately such models present serious computational burdens in estimation.[10] Therefore, as a rule of thumb, practitioners should use richer specifications with constant coefficients and volatilities when the main interest is in point forecasting, and time-varying specifications when the main interest is in density forecasting. Finally, we found that modelling the time variation in both the volatilities and the conditional mean coefficients helps more than only modelling the variation in the volatilities, a result which was robust across different data-sets and is in contrast to other studies that have focused on different target variables (e.g., Carriero, Clark, & Marcellino, 2013b, who focus on nowcasting GDP growth). The time variation in volatilities is particularly useful during recession periods.

5. Conclusions

Previous research (e.g., Favero & Marcellino, 2005) has shown evidence that simple autoregressive models often provide better forecasts of fiscal variables than vector autoregressions. This result is counter-intuitive, because economic theory suggests that fiscal variables should be tightly intertwined, and therefore individual time series should contain useful information about the others.

In this paper we explore the possibility that the VARs considered by previous studies were too small in scale, were probably burdened with overparameterization, and did not feature time variation in the coefficients and volatilities. We estimate several specifications of Bayesian VARs which instead allow the information contained in a large data set to be summarized efficiently, avoid the overparameterization problem, and allow for time variation in both the coefficients and the volatilities.

A second contribution of this paper is to focus on forecasting the whole predictive distribution of fiscal variables, rather than limiting the interest to point forecasts only. This aspect seems particularly relevant for fiscal variables, as assessments of fiscal stability and of the overall credit and default risk of a country should typically be based on the specification of a whole probability distribution for the future state of the economy.

Using data from Germany, France, the UK, and the US, we explore the performances of BVARs with constant and drifting coefficients for forecasting key fiscal variables such as government revenues, expenditure, and interest payments on the outstanding debt. We find that: (i) once overparameterization has been dealt with, the use of additional explanatory variables helps in the forecasting of fiscal variables, and multivariate models outperform univariate specifications in forecasting; (ii) both the adoption of a large system and the introduction of time variation help in forecasting, with the former being more important for point forecasting and the latter for density forecasting; and (iii) the use of drifting coefficients in both the conditional mean parameters and the volatilities provides further help for forecasting relative to models featuring time variation only in the volatilities.

Acknowledgments

We thank the Editor, an associate Editor, and two anonymous referees for several detailed comments and

10 Carriero et al. (2012) and Koop and Korobilis (2013) have each suggested a way around this problem, involving, respectively, the assumption of a factor structure for the drifting volatilities, and the use of an approximation based on forgetting factors.

suggestions. We also thank participants at the CFE and Royal Economic Society conferences. Carriero gratefully aknowledges support for this work from the Economic and Social Research Council [ES/K010611/1].

Appendix A. Marginal likelihood

In order to choose the hyperparameter θ, we maximize the marginal data density $p(Y)$ with respect to such a parameter (and therefore, the data density, when seen as a function of θ, is also referred to as the marginal likelihood of θ). Under the naturally conjugate prior used in this paper, the distributions $p(\Phi, \Sigma|Y)$, $p(Y|\Phi, \Sigma)$, and $p(\Phi, \Sigma)$ are fully known (including the integrating constants), and therefore $p(Y)$ can be obtained simply by inverting the Bayes formula:

$$p_\theta(Y) = \frac{p(Y|\Phi, \Sigma)p_\theta(\Phi, \Sigma)}{p_\theta(\Phi, \Sigma|Y)}. \tag{14}$$

We insert the subscript θ to emphasize the fact that these distributions are conditional on θ (because such a parameter rescales the prior variance matrix Ω_0, see Eq. (3)). The distributions appearing in Eq. (14) are either matricvariate-t or matricvariate-normal (see e.g. Kadiyala & Karlsson, 1997). Moreover, the kernels of the numerator and the denominator of Eq. (14) coincide. It follows that $p_\theta(Y)$ is given by the products and ratios of the integrating constants of $p_\theta(\Phi, \Sigma|Y)$, $p(Y|\Phi, \Sigma)$, and $p_\theta(\Phi, \Sigma)$. Such integrating constants are all available and can be plugged into Eq. (14), which, after some algebra, yields:

$$p(Y) = \pi^{-\frac{TN}{2}} \times |\Omega_0|^{-\frac{N}{2}} \times |\bar{\Omega}|^{\frac{N}{2}} \times |S_0|^{\frac{v_0}{2}} \times |\bar{S}|^{-\frac{v_0+T}{2}} \times \prod_{i=1}^{N} \frac{\Gamma(\frac{v_0+T+1-i}{2})}{\Gamma(\frac{v_0+1-i}{2})},$$

where $\Gamma(\cdot)$ is the univariate gamma function. Noting that $|\Omega_0||\bar{\Omega}^{-1}| = |\Omega_0||X'X + \Omega_0^{-1}| = |X\Omega_0X' + I|$, the expression above resembles the definition of the p.d.f. of a matricvariate t (Dickey, 1967), and coincides with that reported by Carriero et al. (2013a) and Giannone et al. (2012).

Appendix B. MCMC algorithm for the TVP model

Consider the general time-varying VAR model:

$$y_t = \Phi_{c,t} + \Phi_{1,t}y_{t-1} + \Phi_{2,t}y_{t-2} + \cdots + \Phi_{p,t}y_{t-p} + \Sigma_t^{1/2}\varepsilon_t; \quad \varepsilon_t \sim i.i.d.\ N(0, I_N)$$

with:

$$\Phi_t = \Phi_{t-1} + \eta_t, \qquad \mathrm{Var}(\eta_t) = Q$$
$$\mathrm{Var}(v_t) \equiv \Sigma_t = A_t^{-1}H_t(A_t^{-1})',$$

where the structures of A_t and H_t are described in Eq. (7).

B.1. Prior distributions and starting values

The prior for the VAR coefficients (i.e., the initial conditions) is assumed to be normal with mean ϕ_0 and variance v_0. As in the fixed coefficient BVAR, the prior mean equals 0. The prior variance is set using Eq. (3) with the hyperparameter $\theta = 0.2$.[11] Let T_0 denote the length of a training sample. For the US, $T_0 = 40$, while $T_0 = 20$ for the UK and Germany. Let $\hat{v}^{ols}$ denote the OLS estimate of the VAR covariance matrix estimated on the training sample. The prior for the diagonal elements of the VAR covariance matrix (see Eq. (7)) is defined as $\ln h_0 \sim N(\ln \mu_0, I_3)$, where μ_0 are the diagonal elements of the Cholesky decomposition of $\hat{v}^{ols}$. The prior for the off-diagonal elements A_t is $A_0 \sim N\left(\hat{a}^{ols}, V\left(\hat{a}^{ols}\right)\right)$ where $\hat{a}^{ols}$ are the off-diagonal elements of $\hat{v}^{ols}$, with each row being scaled by the corresponding element on the diagonal. $V\left(\hat{a}^{ols}\right)$ is assumed to be diagonal with the elements set equal to 10 times the absolute value of the corresponding element of $\hat{a}^{ols}$.

Let Q_{OLS} denote the OLS estimate of the coefficient covariance matrix using the training sample. Its prior distribution is assumed to be Inverse Wishart, with a scale matrix given by $\bar{Q} = Q_{OLS} \times T_0 \times k$, where the scalar $k = 3.5 \times 10^{-4}$, as per Cogley and Sargent (2005). The prior degrees of freedom are set equal to T_0, the length of the training sample.

The prior distribution for the blocks of S is inverse Wishart: $S_{i,0} \sim IW(\bar{S}_i, p_i)$, where $i = 1, \ldots, 4$ indexes the blocks of S. $\bar{S}_i$ is calibrated using $\hat{a}^{ols}$. Specifically, $\bar{S}_i$ is a diagonal matrix with the relevant elements of $\hat{a}^{ols}$ multiplied by 10^{-3}. Following Cogley and Sargent (2005), we postulate an inverse-gamma distribution for the elements of G, $\sigma_i^2 \sim IG\left(10^{-4}/2, 1/2\right)$.

B.2. Simulating the posterior distributions

The MCMC algorithm is composed of the following steps. We use 20,000 iterations and discard the first 19,000 as a burn-in period.

B.2.1. Drawing the VAR coefficients

The distribution of the time-varying VAR coefficients Φ_t conditional on all other parameters and hyperparameters is linear and Gaussian: $\Phi_t|X_{i,t}, \Xi \sim N\left(\Phi_{T|T}, P_{T|T}\right)$ and $\Phi_t|\Phi_{t+1}, X_{i,t}, \Xi \sim N\left(\Phi_{t|t+1,\phi_{t+1}}, P_{t|t+1,\phi_{t+1}}\right)$, where $t = T-1, \ldots, 1$, Ξ denotes a vector that holds all of the other VAR parameters, and where $\Phi_{T|T} = \mathrm{E}\left(\Phi_T|X_{i,t}, \Xi\right)$, $P_{T|T} = \mathrm{Cov}\left(\Phi_T|X_{i,t}, \Xi\right)$, $\Phi_{t|t+1,\phi_{t+1}} = \mathrm{E}\left(\Phi_t|X_{i,t}, \Xi, \Phi_{t+1}\right)$ and $P_{t|t+1,F_{t+1}} = \mathrm{Cov}\left(\Phi_t|X_{i,t}, \Xi, \Phi_{t+1}\right)$. As was shown by Carter and Kohn (1994), the simulation proceeds by first using the Kalman filter to draw $\Phi_{T|T}$ and $P_{T|T}$, and then proceeding backwards in time using $\Phi_{t|t+1} = \Phi_{t|t} + P_{t|t}P_{t+1|t}^{-1}\left(\Phi_{t+1} - \Phi_t\right)$ and $\Phi_{t|t+1} = \Phi_{t|t} - P_{t|t}P_{t+1|t}^{-1}P_{t|t}$.

B.2.2. Drawing the elements of H_t

Following Cogley and Sargent (2005), the diagonal elements of the VAR covariance matrix are sampled

[11] Unlike in the case of the fixed coefficient VAR, a closed form for the marginal likelihood is not available for the TVP-VAR, and MCMC methods are required. Given that the TVP model is estimated recursively, computing and maximising the marginal likelihood for each recursion entails an extremely high computational burden.

using the Metropolis Hastings algorithm of Jacquier, Polson, and Rossi (1994). Given a draw for Φ_t, the VAR model can be written as $A_t' v_t = u_t$, where $v_t = y_t - \Sigma_{l=1}^{p} \Phi_{l,t} y_{t-l}$ and $\mathrm{Var}(u_t) = H_t$. Jacquier et al. (1994) note that, conditional on other VAR parameters, the distribution h_{it}, $i = 1, \ldots, 3$, is given by $f(h_{it}|h_{it-1}, h_{it+1}, u_{it}) = f(u_{it}|h_{it}) \times f(h_{it}|h_{it-1}) \times f(h_{it+1}|h_{it}) = h_{it}^{-0.5} \exp\left(\frac{-u_{it}^2}{2h_{it}}\right) \times h_{it}^{-1} \exp\left(\frac{-(\ln h_{it}-\mu)^2}{2\sigma_{h_i}}\right)$, where μ and σ_{h_i} denote the mean and variance of the log-normal density $h_{it}^{-1} \exp\left(\frac{-(\ln h_{it}-\mu)^2}{2\sigma_{h_i}}\right)$. Jacquier et al. (1994) suggest using $h_{it}^{-1} \exp\left(\frac{-(\ln h_{it}-\mu)^2}{2\sigma_{h_i}}\right)$ as the candidate generating density, with the acceptance probability defined as the ratio of the conditional likelihood $h_{it}^{-0.5} \exp\left(\frac{-u_{it}^2}{2h_{it}}\right)$ at the old and new draws. This algorithm is applied at each period in the sample.

B.2.3. Drawing the elements of A_t

Given a draw for Φ_t, the VAR model can be written as $A_t' v_t = u_t$, where $v_t = y_t - \Sigma_{l=1}^{p} \Phi_{l,t} y_{t-l}$ and $\mathrm{Var}(u_t) = H_t$. This is a system of equations with time-varying coefficients, and, given a block diagonal form for $\mathrm{Var}(\tau_t)$, the standard methods for state space models described by Carter and Kohn (1994) can be applied.

B.2.4. Drawing the hyperparameters

Conditional on y_t, Φ_t, H_t, and A_t, the innovations to Φ_t, H_t, and A_t are observable, which allows us to draw the hyperparameters – the elements of Q, S, and σ_i^2 – from their respective distributions.

References

Amisano, G., & Giacomini, R. (2007). Comparing density forecasts via weighted likelihood ratio tests. *Journal of Business and Economic Statistics*, *25*, 177–190.

Arghyrou, M., & Kontonikas, A. (2012). The EMU sovereign-debt crisis: Fundamentals, expectations and contagion. *Journal of International Financial Markets, Institutions and Money*, *22*(4), 658–677.

Artis, M. J., & Marcellino, M. (2001). Fiscal forecasting: The track record of the IMF, OECD and EC. *Econometrics Journal*, *4*(1), 20–36.

Banbura, M., Giannone, D., & Reichlin, L. (2010). Large Bayesian vector autoregressions. *Journal of Applied Econometrics*, *25*(1), 71–92.

Blanchard, O., & Perotti, R. (2002). An empirical characterization of the dynamic effects of changes in government spending and taxes on output. *Quarterly Journal of Economics*, *117*(4), 1329–1368.

Breeden, D. T. (2012). *Consumption as a leading indicator*. Working paper, Massachusetts Institute of Technology and Duke University, April.

Carriero, A., Clark, T., & Marcellino, M. (2012). *Common drifting volatility in large Bayesian VARs*. CEPR Discussion Papers 8894, C.E.P.R. Discussion Papers.

Carriero, A., Clark, T., & Marcellino, M. (2013a). *Bayesian VARs: specification choices and forecast accuracy*. CEPR Discussion Papers 8273, C.E.P.R. Discussion Papers.

Carriero, A., Clark, T., & Marcellino, M. (2013b). *Real-time nowcasting with a Bayesian mixed frequency model with stochastic volatility*. CEPR Discussion Papers 9312, C.E.P.R. Discussion Papers.

Carriero, A., Kapetanios, G., & Marcellino, M. (2011). Forecasting large datasets with Bayesian reduced rank multivariate models. *Journal of Applied Econometrics*, *26*(5), 735–761.

Carter, C. K., & Kohn, R. (1994). On Gibbs sampling for state space models. *Biometrika*, *81*(3), 541–553.

Cassidy, G., Kamlet, M. S., & Nagin, D. S. (1989). An empirical examination of bias in revenue forecasts by state governments. *International Journal of Forecasting*, *5*(3), 321–331.

Clark, T. (2011). Real-time density forecasts from BVARs with stochastic volatility. *Journal of Business and Economic Statistics*, *29*(3), 327–341.

Clark, T., & McCracken, M. (2013). Advances in forecast evaluation. In G. Elliott & A. Timmermann (Eds.), *Handbook of economic forecasting, Vol. 2 (B)* (pp. 1107–1201). Elsevier.

Cogley, T., Morozov, S., & Sargent, T. J. (2005). Bayesian fan charts for UK inflation: forecasting and sources of uncertainty in an evolving monetary system. *Journal of Economic Dynamics and Control*, *29*(11), 1893–1925.

Cogley, T., & Sargent, T. J. (2005). Drifts and volatilities: monetary policies and outcomes in the post WWII US. *Review of Economic Dynamics*, *8*(2), 262–302.

D'Agostino, A., Gambetti, L., & Giannone, D. (2013). Macroeconomic forecasting and structural change. *Journal of Applied Econometrics*, *28*(1), 82–101.

Dickey, J. (1967). Matricvariate generalizations of the multivariate *t* distribution and the inverted multivariate *t* distribution. *Annals of Mathematical Statistics*, *38*(2), 511–518.

Diebold, F. X., & Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics*, *13*(3), 253–263.

Doan, T., Litterman, R., & Sims, C. (1984). Forecasting and conditional projection using realistic prior distributions. *Econometric Reviews*, *3*(1), 1–100.

Favero, C. A., & Marcellino, M. (2005). Modelling and forecasting fiscal variables for the euro area. *Oxford Bulletin of Economics and Statistics*, *67*(1), 755–783.

Feenberg, D. R., Gentry, W., Gilroy, D., & Rosen, H. S. (1989). Testing the rationality of state revenue forecasts. *Review of Economics and Statistics*, *71*(2), 300–308.

Geweke, J., & Amisano, G. (2010). Comparing and evaluating Bayesian predictive distributions of asset returns. *International Journal of Forecasting*, *26*(2), 216–230.

Giannone, D., Lenza, M., & Primiceri, G. (2012). *Prior selection for vector autoregressions*. WP ECARES 2012-002, Universite Libre de Bruxelles.

Harvey, D., Leybourne, S., & Newbold, P. (1997). Testing the equality of prediction mean squared errors. *International Journal of Forecasting*, *13*(2), 281–291.

Jacquier, E., Polson, N., & Rossi, P. (1994). Bayesian analysis of stochastic volatility models. *Journal of Business and Economic Statistics*, *12*(4), 371–418.

Jonung, L., & Larch, M. (2006). Improving fiscal policy in the E.U.: the case for independent forecasts. *Economic Policy*, *21*(47), 491–534.

Kadiyala, K., & Karlsson, S. (1997). Numerical methods for estimation and inference in Bayesian VAR-models. *Journal of Applied Econometrics*, *12*(2), 99–132.

Koop, G. (2013). Forecasting with medium and large Bayesian VARs. *Journal of Applied Econometrics*, *28*(2), 177–203.

Koop, G., & Korobilis, D. (2013). Large time-varying parameter VARs. *Journal of Econometrics*, *177*(2), 185–198.

Leal, T., Pérez, J., Tujula, M., & Vidal, J. P. (2008). Fiscal forecasting: lessons from the literature and challenges. *Fiscal Studies*, *29*(3), 347–386.

Leeper, E. M., Sims, C. A., & Zha, T. (1996). What does monetary policy do? *Brookings Papers on Economic Activity*, *2*(1), 1–78.

Litterman, R. (1986). Forecasting with Bayesian vector autoregressions—five years of experience. *Journal of Business and Economic Statistics*, *4*(1), 25–38.

Marcellino, M. (2006). Some stylized facts on non-systematic fiscal policy in the Euro area. *Journal of Macroeconomics*, *28*(3), 461–479.

Marcellino, M., Stock, J., & Watson, M. (2006). A comparison of direct and iterated multistep AR methods for forecasting macroeconomic time series. *Journal of Econometrics*, *127*(1–2), 499–526.

Melander, A., Sismanidis, G., & Grenouilleau, D. (2007). *The track record of the Commission's forecasts — an update*. European Economy, Economic Papers 291, DG ECFIN, European Commission, Brussels.

Primiceri, G. (2005). Time varying structural vector autoregressions and monetary policy. *The Review of Economic Studies*, *72*(3), 821–852.

Robertson, J. C., & Tallman, E. W. (1999). *Prior parameter uncertainty: Some implications for forecasting and policy analysis with VAR models*. Working paper 99-13, Federal Reserve Bank of Atlanta.

Sims, C. A., & Zha, T. (1998). Bayesian methods for dynamic multivariate models. *International Economic Review*, *39*(4), 949–968.

Stock, J. H., & Watson, M. W. (2002). Macroeconomic forecasting using diffusion indexes. *Journal of Business and Economic Statistics*, *20*(2), 147–162.

Strauch, R., Hallerberg, M., & von Hagen, J. (2004) *Budgetary forecasts in Europe — the track record of stability and convergence programmes*. European Central Bank, WP 307.

von Hagen, J., Schuknecht, L., & Wolswijk, G. (2011). Government bond risk premiums in the E.U. revisited: the impact of the financial crisis. *European Journal of Political Economy*, *27*(1), 36–43.

Waggoner, D. F., & Zha, T. (1999). Conditional forecasts in dynamic multivariate models. *The Review of Economics and Statistics*, *81*(4), 639–651.

Zellner, A. (1971). *An introduction to Bayesian inference in econometrics*. Wiley.

Zha, T. (1998). A dynamic multivariate model for use in formulating policy. *Economic Review*, *1*, 16–29.